Hanging out in a bad neighbourhood:
Does the perceived credibility of online business directories 
affect user attitudes toward the services they represent.

A thesis presented in partial fulfillment of the requirements for the degree of

Master of Information Systems

at Massey University, Albany, New Zealand.

Daron Robinson
2011
Abstract

The effect of aesthetics on the overall credibility of a website is reasonably well studied, but there is little understanding of the degree to which the user’s perception of the credibility of a website may result in a transference of attitude to the subject of the website, particularly towards the people, products or services it may represent. This issue is particularly unexplored with regard to web business directories that mediate between users and services and third party advertisers in a Multi Sided Online Market.

This research studies the impact of different qualities of aesthetic appeal, content relevance and content reputability on the user’s affect and their perception of the credibility of services represented by such intermediary websites.

Four variations of a fictional web business directory were employed, three of which reduce either the visual appeal or relevance and reputability of content compared to a control version. This website was used by participants in the study to assess the credibility of services depicted on the website while each of the versions was displayed, and to gather qualitative comments.

The results showed a significant impact on user’s impressions of the credibility of services shown, especially in the case of the presence of irrelevant advertising, but also in the case of an unattractive intermediary website, demonstrating that users do not consider the services truly distinct from their context, but see them as part of a wider unified social system. Users were found to be judging the implied credibility of all the participants in a Multi Sided Online Market by their associations with each other and also by the aesthetic condition of the marketplace they are mutually inhabiting. This behaviour mirrors similar credibility assessments seen in true social contexts like a physical neighbourhood.

The implications of this result for the designers and owners of Advertising supported and ‘Freemium’ Online Directories is discussed and directions for future credibility and aesthetic research proposed.
Acknowledgements

I sincerely thank my research supervisor David Paulson for his continued support and guidance over this long campaign.

Thanks also to Darry Forsyth for his invaluable help with my research methodologies, and for recommending the rather excellent book “Discovering Statistics using SPPS (and sex and drugs and rock 'n' roll)”.  

Thanks to Jodie & John and to Al & Mandy who kindly let me use their houses to get some space alone to concentrate. And also to Shane, from GrabOne.co.nz, who gave me time off to work on my thesis even when there was really no time to spare.

To Jim and Val, who generously assisted us when study was soaking up all my time and eating into our families income, I also offer my humble gratitude.

But above all I thank my wife Alexandra and our children who encouraged and supported me at every step. Without their patience and sacrifice this work would not have been possible.
# Table of Contents

1 Introduction ........................................ 11

2 Literature Review ................................... 14
   2.1 User Experiences and Attitudes ................. 14
      2.1.1 Human-Computer Interaction .............. 14
      2.1.2 Defining User Affect ...................... 15
   2.2 Aesthetics and Affect .......................... 17
      2.2.1 Aesthetic Models .......................... 17
      2.2.2 Elements of Aesthetics on the Web ...... 18
      2.2.3 Aesthetics and Affect .................... 20
   2.3 Mixing the Message ............................. 21
      2.3.1 The impact of Advertising ................ 21
      2.3.2 Advert Avoidance .......................... 22
      2.3.3 Irrelevant Added Content ................ 23
   2.4 The Perception of Credibility .................. 25
      2.4.1 Credibility and Trust ..................... 25
      2.4.2 Aesthetics and Credibility ............... 26
      2.4.3 Credibility and Online Personas .......... 26
      2.4.4 Reputation Systems ....................... 29
      2.4.5 Warranting .................................. 29
      2.4.6 Symptoms of Trustworthiness .............. 30
      2.4.7 The Anthropomorphic Web ................ 31
   2.5 A Multi Sided Exchange .......................... 33
      2.5.1 The Two Sided Market Model ............... 33
      2.5.2 Online Business Directories .............. 35
      2.5.3 A Love - Hate relationship ............... 36
   2.6 A Multi Sided Network of Affect ................ 39
   2.7 Bad Neighbourhoods on the Web ................. 41

3 Research Design and Methodology ................. 42
   3.1 Research Methodology Overview ................ 42
      3.1.1 Research Website and Instrument Design .... 42
      3.1.2 Pilot Study ................................ 43


4 Validating the Designs – The Pilot Study

4.1 Pilot Study Design
  4.1.1 Study of Design Features and Ranking
  4.1.2 Study of Advertising Ranking

4.2 Pilot Study Participation

4.3 Comparative Website Design Results
  4.3.1 Semantic Differential Question Results
  4.3.2 Freeform comments
  4.3.3 Manual Ranking Results
  4.3.4 Conclusions of the Website Design Component

4.4 Comparative Advertising Results
  4.4.1 Conclusions of the Advertising Component

4.5 Conclusions and modifications to the Website Designs

5 Between Groups Research
5.1 Between Groups Research Overview
5.2 Between Groups Instrument Design
  5.2.1 Simulating an Online Directory
  5.2.2 Functionality of the Website
5.3 Between Groups Research Results
  5.3.1 Participants & Recruitment
5.4 Analysis
  5.4.1 Validating the data
  5.4.2 Validating the survey questions and extracting factors
  5.4.3 Analysis with Two-Way Independent Factorial Anova
5.5 Limitations of this study

6 Within Subjects Research
  6.1 Within Subjects Research Objectives
  6.2 Within Subjects Research Instrument
    6.2.1 Counterbalancing
  6.3 Within Subjects Research Results
    6.3.1 Participants & Recruitment
  6.4 Analysis
    6.4.1 Validating the Data
    6.4.2 Validating the survey questions and extracting factors
    6.4.3 Analysis with Two-Way Independent Repeated Measures Anova

7 Qualitative Analysis
  7.1 Nature of Analysis
  7.2 Coding of Comments
  7.3 Coding Results
    7.3.1 Overall Results
    7.3.2 Results for Between Groups study
    7.3.3 Results for Within Subjects study

8 Discussion
  8.1 Introduction
  8.2 Recapitulation
  8.3 Contribution
    8.3.1 Implications for Multi-Sided Online Platforms
    8.3.2 Implications for Affect & Credibility Research
    8.3.3 Implications for Web Design Practice
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Overall Limitations &amp; Future Research</td>
<td>113</td>
</tr>
<tr>
<td>9.1</td>
<td>Limitations</td>
<td>113</td>
</tr>
<tr>
<td>9.2</td>
<td>Future Research</td>
<td>115</td>
</tr>
<tr>
<td>10</td>
<td>Conclusion</td>
<td>116</td>
</tr>
<tr>
<td>11</td>
<td>References</td>
<td>118</td>
</tr>
<tr>
<td>Appendix 1</td>
<td>Design Source Websites</td>
<td>125</td>
</tr>
<tr>
<td>Appendix 2</td>
<td>Pilot Study</td>
<td>132</td>
</tr>
<tr>
<td>Appendix 3</td>
<td>Information, Consent and Survey Forms</td>
<td>134</td>
</tr>
</tbody>
</table>
List of Illustrations:

Illustration 1: HSV Colour Cone 19
Illustration 2: Social-Technical Gap (Whitworth, 2006) 28
Illustration 3: Two Sided Market Model example 34
Illustration 4: Multi Sided Online Market 37
Illustration 5: Multi Sided Network of Affect 39
Illustration 6: Pilot Screenshots 46
Illustration 7: Colour use in the Attractive design 47
Illustration 8: Colour use in the Unattractive design 48
Illustration 9: Contrast use in the Attractive Design 49
Illustration 10: Contrast use in the Unattractive design 49
Illustration 11: Unbalanced shapes in the Unattractive design 50
Illustration 12: Dimensional effects in the Attractive design 51
Illustration 13: Background textures in the Unattractive design 51
Illustration 14: Padding and linespacing in the Unattractive design 52
Illustration 15: Random positioning in the Unattractive Design 53
Illustration 16: Consistent Fonts & Colours in the Attractive Design 53
Illustration 17: Inconsistent Fonts & Colours in the Unattractive Design 54
Illustration 18: Animation in the Unattractive Design 55
List of Tables:

Table 1: Survey Questions 66

Table 2: Pilot Study Questions 69

Table 3: Pilot Study results for Terrible - Wonderful question (ranked by mean) 72

Table 4: Pilot Study results for Dull - Stimulating question (ranked by mean) 72

Table 5: Pilot Study results for Ugly - Beautiful question (ranked by mean) 73

Table 6: Pilot Study results website * Direction of Comment cross tabulation 74

Table 7: Pilot study results - web page design ranking 75

Table 8: Pilot Study results - Advertisement Ranking 76

Table 9: Skewness and Kurtosis results 83

Table 10: Two way Factorial ANOVA - tests of Between Groups effects 85

Table 11: Skewness and Kurtosis results 91

Table 12: Two way Repeated Measures ANOVA - tests of Within Subjects effects 93

Table 13: Within Subjects Contrasts 93

Table 14: Themes used in coding 96

Table 15: Overall coding results 97

Table 16: Between Groups coding results 98

Table 17: Within Subjects coding results 99
1 Introduction

In the past, when people used websites, one might perhaps have assumed that their intent was mostly to be informed, to execute a simple task, or perhaps occasionally to be entertained. Users were presumed to conceive of the website as a neutral tool (Fishwick, 2006) that existed in an abstract realm called the world wide web, a universal source of knowledge and information which they could read and interact with to achieve their goals, the whole exercise being somehow detached from the fickle world of human emotion and social interaction.

Certainly, historically, this dry mechanistic vision is largely how Human Computer Interaction was studied and defined (Bevan, Kirakowski, & Maissel, 1991). But the reality, especially in this day of social networking, has become more complex, as Winter, Saunders and Hart (2003) depict it, “websites are on-stage work areas where a performance is given to an actual or implied audience of potential customers, employees, suppliers, partners, and regulators” (p.311). That is a cast of characters to which one might now add: potential lovers, friends, workmates, pupils, followers and co-conspirators.

The humble website then, has grown to become a pervasive platform for expression and communication that embraces the whole eclectic range of human thought, knowledge and feeling and presents it all equally to anyone with access to a web browser (Nardi, Schiano, Gumbrecht, & Swartz, 2004),(D. M. Boyd & Ellison, 2007).

It is no surprise therefore, that more recent studies (Porat & Tractinsky, 2008) have begun to reveal that emotion, or Affect, is unavoidably involved at both ends of the website interaction, both in the intent of the designer and the experience of the user (Beale & Peter, 2008). Although the computer mediates the interaction, it is always people who first devise and then interpret the message.

Increasingly, the impact of the perceived aesthetics of websites has been noted (Lavie & Tractinsky, 2004) and this purely visual aspect has been shown to have a surprisingly rapid and powerful influence (Lindgaard, Fernandes, Dudek, & J. Brown, 2006) on user emotions and attitudes, especially toward the credibility of the website.

Similarly, user perceptions of the relevance and reputability of associated advertising has been seen to sway user's emotional response to websites (Fogg et al., 2001), often without any conscious recognition that the advertising was even present (Burke, Gorman, Nilsen, & Hornof, 2004).
This apparently inevitable surge of emotional responses is often directed at the website as if it were human, a social actor with whom the user is engaging. Norman (2005) describes this anthropomorphic effect, claiming that despite the fact that the website is not a “causal agent” the interacting user will “build up expectations of behaviour based on prior experience, and if the items with which you interact fail to live up to expectations, that is a violation of trust, for which you assign blame, which can soon lead to anger” (p.139).

Just like a newly met stranger, it seems websites will often end up being viewed with extreme caution, or trusted implicitly, all on the strength of swift, sketchy interpretations from a few available socially relevant and largely visual cues (Riegelsberger, Sasse, & McCarthy, 2004).

While this rush to judgement is perhaps unavoidable, just another side effect of the uncertainties of any human exchange, the issue becomes more clouded when the website that is inspiring confidence or distrust is not itself an embodiment of one of the social actors in the exchange, but is instead a Directory, no more than a spokesman, an intermediary, a messenger bearing news upon the virtual stage. And like Shakespeare’s Cleopatra, are we tending to judge rashly, based only on the appearance and affect of the messenger, fearing deceit where there is none:

“But there’s no goodness in thy face: if Antony
Be free and healthful,—so tart a favour
To trumpet such good tidings! If not well,
Thou shouldst come like a Fury crown’d with snakes,
Not like a formal man” (Antony and Cleopatra, Act 2, Scene 5).

It is uncertain how and why users form their impressions about services seen on online directories. Do user’s tend to “shoot the messenger” for the content of message, or instead despise the message due to the appearance of the messenger. Do those services, products or personas depicted on the directory become guilty by association, or perhaps shine more brightly by being well presented and in good company, being seen in a clean well-lighted place as Hemingway would have it?

It is this gap in understanding that the study will address by examining how the perceived credibility of services and personas are affected when presented on web directory web pages of reduced visual appeal or reduced relevance and reputability compared to a control version of the website.

Introduction
This study will review the existing literature on the effect of website aesthetics and advertising relevance and reputability on the formation of credibility assessments and also examine how multi-sided market mediators like Online Business Directories expose users to mixed messages about the services they are representing.

Two experiments will be described that examine the effect of variations in Aesthetics and Advertising relevance on users impressions of the services presented on an Online Directory, using both a between groups and a within subjects design.

The results of these experiments will be analysed and discussed in terms of the wider sociotechnical effect these Online Directory Services may be having and the consequent implications for business promotion and advertising.
2 Literature Review

2.1 User Experiences and Attitudes

2.1.1 Human-Computer Interaction

The field of Human Computer Interaction has in recent years begun to embrace a broader understanding of what truly constitutes a meaningful interaction between users and machines.

Historically, HCI research focused mainly on aspects of interface usability on standalone computer systems (Bevan et al., 1991), perhaps most famously through the extensive research of Jakob Nielsen (1994). The arrival of the world wide web quickly demonstrated the need for more rigour in the usability aspects of web page design (Bevan, 1997), (Nielsen, Tahir, & Tahir, 2001), but essentially this earlier work was largely confined to the study of interfaces as tools rather than as media (Fishwick, 2006), neglecting, some would say deliberately (Tractinsky, 2004), (Hassenzahl, 2004) issues of form over function, or emotion over reason. According to Hassenzahl (2004) “HCI was, at best, suspicious about beauty. 'If it is pretty, it won’t work,’ summarises one of the common prejudices” (p.320).

But in recent years, with the transition of static websites into richer, more media-like experiences (Fishwick, 2006), a broad trend in research has developed that has begun to embrace issues of user mood as well as user action, and attempts to define how the two impact on one another.

This growth in study of what has come to be termed user affect emerged from very tentative exploratory works in the early days of the web (Kurosu & Kashimura, 1995), (Tractinsky, 1997). Kim and Moon (1998) in their seminal study on Emotional Usability described computer interaction as having similar emotional impact on users as watching a film, declaring that “computer interfaces can basically be considered as a collection of snap shots that are composed of diverse visual and auditory stimuli designed to directly engage the user into the interaction” (p.4).

This early exploration of user affect has since been expanded and enhanced by a number of empirical studies (Jordan, 1998), (Hu et al., 2004), (Zhang & N. Li, 2005) that attempt to measure and to some extent predict and control (Egger, 2001), (Porat & Tractinsky, 2008) the ways user's affect might be modified by participating in various computer interactions, either in the process of directly controlling software as considered by Beale and Peter (2008), in the broader realms of
computer mediated communication (CMC) between users and organisations (Hampton-Sosa & Koufaris, 2005), or even directly with other users in scenarios as loaded with complex human emotion as Ellison, Heino, and Gibbs (2006) Online Dating scenarios.

2.1.2 Defining User Affect

The concept of affect is defined by Norman (2005) as “the general term for the judgemental system, whether conscious or subconscious” (p.11). He goes on to relate it to emotion, which “is the conscious experience of affect, complete with attribution of its cause and identification of its object” (p.11).

The affect of website users has been shown to be influenced by multiple aspects of the processes, systems and interactions they are experiencing. Rafaeli & Vilnai-Yavetz (2004) identify three aspects of affective response to events and objects:

Aesthetics - which they express as “the sensory experience it elicits, and the extent to which this experience fits individual goals and spirit” (p.95).

Instrumentality - which they define as “the extent to which the artefact contributes to performance or to promoting goals” (p.94). They equate this concept of instrumentality with usability as defined by the HCI community.

Symbolism - “the meanings or associations” (p.95) an artefact elicits. They point out that symbolic meanings attributed to an artefact are determined by the users interpretation and context, and may have little to do with the intent of the originator of the artefact.

These three concepts of affective response can be equated to Normans (2005) concepts of Emotional design:

Visceral Design, which “concerns itself with appearances” (p.5).

Behavioural design which has to do “with the pleasure and effectiveness of use” (p.5).

Reflective design which “considers the rationalisation and intellectualisation of a product” (p.5).

Within the context of the web, Visceral Design can be seen as largely an aspect of the Aesthetic appeal of the website, Behavioural Design is related to Instrumentality via the broad issue of website usability, and Reflective Design is associated with the Symbolism or meaning of the...
website content. These three each play a role in building the users perception of what the site is, does, and means.

Norman (2005) describes the combination of these perceived features as the “system image” (p.76), the sum total of information actually communicated by the website, and each of these aspects inevitably triggers a response in the user, whether intended or not by the original creator of the website.

The experience of using a website will therefore have either a positive or negative effect on the user’s affect, but negative experiences tend to have more impact. Drawing on earlier studies into interpersonal affective responses (K. Edwards & Von Hippel, 1995), Norman (2005) shows that “the affective system makes judgements and quickly helps you determine which things in the environment are dangerous or safe, good or bad” (p.11). While Norman concluded that “happy people are more effective in finding alternative solutions and as a result, are tolerant of minor difficulties” (p.20), overall there has been shown to be a strong tendency for negative bias toward stimuli (D. L. Hamilton & Zanna, 1972), (Coover & Reeder, 1990), (Forgas, 2006) Baumeister, Bratslavsky, Finkenauer and Vohs (2001) went as far as to declare that it is a “basic, pervasive fact of psychology that bad is stronger than good” (p.360).

Precisely which particular aspects of a user’s diverse interactions on the web have the most impact on affect, both good or bad, is an area of ongoing research (Sicilia & Ruiz, 2007), (Beale & Peter, 2008), (Lim et al., 2008), but what has become clear is that the user’s affect is inescapably modified by their experience of using a website, and that this process is largely unconscious and can be extremely rapid.

Lindgaard et al (2006) showed that users form a first impression of a website within 50 milliseconds, based purely on emotional responses to visual stimuli, and that this first impression is enduring, as “even if a website is highly usable and provides very useful information presented in a logical arrangement, this may fail to impress a user whose first impression of a site was negative” (p.115).

This visual first impression effect has been confirmed by others (Tractinsky, Cokhavi, Kirschbaum, & Sharfi, 2006), and suggests that whether website creators are aware of it or not, users are inevitably taking away affective, emotional impressions of their sites that may or may not be what they intended to communicate.
2.2 Aesthetics and Affect

2.2.1 Aesthetic Models

Some may suggest that aesthetics are too subjective a realm to measure or model, but in that sense the study of aesthetics is very similar to the study of affect, for as Karvonen (2000) maintains:

Aesthetics and emotions share a common quality: they have often been mystified, and this is why it has been claimed that they cannot be really studied, at least not in the same way as more rational things can be studied. Furthermore, aesthetics, if not emotions, have often been considered to be unique - everyone has a taste of their own, so universal beauty assessments do not hold. However, people might not be so independent in their preferences and likes, but often follow some general principles of "styles" or "trends" or "fashions", in what they consider to be beautiful. (p.87)

Certainly, the particular importance of website aesthetics in interaction design and usability has been well studied in recent years. Tractinsky, Katz and Ikar (2000) claimed that what is beautiful is usable and despite some contrary findings (Hassenzahl, 2004) considerable research has established the deeper significance of many aspects of aesthetics in interaction design (Karvonen, 2000) (Lavie & Tractinsky, 2004), (Rafaeli & Vilnai-Yavetz, 2004), (Tractinsky & Zmiri, 2006), (Löwgren, 2009).

Tractinsky (2004) presented a model of aesthetics drawn from the classical notions of architecture expressed by Vitruvius, which consisted of *firmitas* (strength), *utilitas* (utility), and *venustas* (beauty). He maintained that it is from the marriage of form and function implied by these architectural principles that aesthetics within the realm of HCI have descended.

Tractinsky proposed that users' aesthetic perceptions consist of two primary dimensions, ‘classical aesthetics’ and ‘expressive aesthetics’.

**Classical aesthetics** correlates to aesthetic principles that focus on order and clarity in design, similar to many of the design rules advocated by usability experts like Nielsen.

**Expressive aesthetics** is demonstrated in the creative or unique aspects of the design, by the ways the design branches from convention or rules.

Follow up studies showed “that classical aesthetics is strongly correlated with evaluations of usability; that both aesthetic dimensions contributed to user satisfaction, and that expressive
aesthetics is a better determinant of immediate attractiveness impressions of Websites” (Porat & Tractinsky, 2008, p. 146).

2.2.2 Elements of Aesthetics on the Web

A successful and well designed web page ideally needs to communicate multiple simultaneous integrated messages, neatly prioritised for consumption, each separately pleasing and yet somehow forming a whole that is attractive and compelling. It is an aesthetic balancing act over which both cultural influences (Tractinsky, 1997), and individual tastes hold significant sway.

There are numerous variations in aesthetic style on the web, some obviously quite different but arguably equally effective in terms of popularity (Ivory & Hearst, 2002), and although there is no fixed best solution in this subjective realm, some patterns and anti-patterns of presentation have evolved over time (Ivory & Megraw, 2005). Recurring concepts around clarity, balance, colour and contrast keep emerging from the literature (Zettl, 1973), (Hoffmann & Krauss, 2004), (Everard & Galletta, 2006), (Moshagen & Thielsch, 2010) although often appearing as synonyms and overlapping terms as new models are proposed.

But one very major theme observed in the literature, especially for larger web pages, is the general problem of Complexity (Michailidou, Harper, & Bechhofer, 2008).

As the web has matured, and increasing amounts of information have begun to be distributed through it, issues of visual and cognitive complexity have arisen that cannot simply take their design cues from solutions in static print media (Harper, Michailidou, & Stevens, 2009). Although nominally a scrolling web page has an unlimited space available to present information, the practical boundaries are set by human perception and attention span rather than technical limits (Tuch, Bargas-Avila, Opwis, & Wilhelm, 2009).

The more information of various sorts that is competing for a users attention and time, the more likely that conflicts and mismatches will occur that diminish the user's perception of a website and their ability to absorb the information on offer. Michailidou et al (2008) in particular found that “visual complexity of a page is negatively related with user perception of how organised, clear, clean and beautiful a page looks” (p.222).

Finding exact metrics for website complexity is not straightforward, because of the subjectivity of personal taste and cultural biases. More objective methods such as counting occurrences of certain aspects of a design like font choices or colour usage have been used with some limited success (Ivory & Megraw, 2005) but the ongoing trends and fashions in web page design make it a shifting target for longitudinal studies. Tuch et al (2009) went to the extent of deriving a metric

Literature Review
for visual complexity based on software compression algorithms, maintaining that “the size of compressed image files predicts the subjective complexity of images generated by websites” (p. 704). Regardless of the difficulties with exact metrics, some of the most important aesthetic features relating to complexity in web page design can at least be summarised:

**Colour & Contrast:** Colour and contrast are both optical effects derived from the intensity and wavelength of light rather than depending on shape or form, and are often combined into one of various popular Colour Theory models such as the HSV Colour Wheel (Hue, Saturation & Value) where any possible combination can be modelled as hue (colour), saturation (how rich the colour is), and value (how dark the colour is). This model is often represented as a wheel, cylinder or cone to depict all three dimensions.

![Illustration 1: HSV Colour Cone](image)

Reflected Light, in combinations of various different colours and tones is the fundamental means by which all visual information is communicated to the eye, so increasing the number, or range of colours in a given area requires greater cognitive resources to interpret, adding to the perceived complexity. Accordingly, users find that web pages that display too many colours or contrasts “tire and confuse them, and make them spend more time on the page than necessary” (Harper et al., 2009, p. 14).

**Shape & Texture:** Shape and texture concern the discrete volume that a visual object occupies, and the nature of the edges, implied surfaces and suggested dimensions. Texture concerns the discourse of tones and transitions perceived to be within an object, and Shape is the distinction between an object and its background. The presence of too many distinct shapes and textures within a space will compete for attention and add clutter and complexity, for instance in the common case of a textured background making the foreground text illegible (Everard & Galletta, 2006).

**Positioning & Composition:** These aspects are concerned with how objects are seen next to each other, their alignments, the spaces created, orientations, rotations, the implied relationships
between things in space. Composition is the broader whole, it can be considered the sum of all the objects and their proximities to one another, the patterns and groupings that they form across the entire visible frame.

Lack of concern with the effects of these relationships adds complexity by creating unintended points of connection and juxtapositions, unclear relationships, conflicts and imbalances that undermine the whole. Everard & Galletta (2006) contend with regard to alignment, that “crooked columns and inconsistent word and line spacing can make it difficult to follow the text horizontally and vertically” (p. 32). Even a few things, balanced badly, can be much worse than many things arranged well (Michailidou et al., 2008).

**Consistency & Flow:** This effect of consistency and flow dictates the movement of the eye over the total design (Sicilia & Ruiz, 2007), the priority and emphasis given to various items over others, and the ability to group and consider multiple items by their similarities and differences in proximity, pattern, shape & texture, colour or contrast (Michailidou et al., 2008).

If inconsistent but similar items appear to be given equal emphasis, or if items are close but not exactly in alignment with one another (Everard & Galletta, 2006), the eye is drawn to these anomalies, complicating the sense of flow and meaning of the whole. Moshagen and Thielsch (2010) propose that “the more fluently a perceiver can process an object, the more positive is his or her aesthetic response” (p.5).

### 2.2.3 Aesthetics and Affect

Determining practical, reproducible methods for designing websites to specifically influence user affect using aesthetic principles has proven an area of some debate, as there are divergent theories (Hassenzahl, 2004) and even fewer solid experimental metrics (Hoffmann & Krauss, 2004), (Tractinsky & Lowengart, 2007) for measuring loose subjective concepts such as beauty or appeal than there are for complexity.

Some studies stress the importance of the interaction *gestalt* (Fishwick, 2006), (Lim, Stolterman, Jung, & Donaldson, 2007), (Chandler, 2007), the total simultaneous perception of the interaction from the users perspective, which “suggests that higher order (“tertiary” or “emergent”) qualities are seen directly in things when we look at them as a whole, but do not apply to atomic parts into which things can be decomposed” (Lavie & Tractinsky, 2004).

Other studies (Park, Choi, & J. Kim, 2004), (Coursaris, Sweirenga, & Watrall, 2008), (Michailidou et al., 2008) take the position that the component parts of a design can be varied and analysed to measure the changes they make in user affect.
Despite the gulf between these various divergent positions, the common thread across many findings indicates that user affect toward websites is being consistently and powerfully impacted by the aesthetics users encounter, and importantly, this is often above any other factor like usability or even content usefulness (Lavie & Tractinsky, 2004).

The potency of aesthetics in altering user affect extends beyond first impressions. De Angeli, Sutcliffe and Hartmann (2006) contend that “not only is beauty an important quality of a product but its effect seems to transcend the object and influence other judgements, in what is known as the halo effect” (p.2710). Related studies (Hartmann, Sutcliffe, & De Angeli, 2007) went on to suggest that users’ overall impression of a website could be a determinant of user satisfaction and system acceptability, even outweighing a poor usability experience or poor content.

Similarly, Tractinsky et al (2000) maintain that “the positive affect created by aesthetically appealing interfaces may be instrumental in improving users’ performance as well” (p.141).

Importantly however, the power of negative affect from an unappealing website experience is even greater, as Lindgaard et al (2006) point out, “a confirmation bias occurring in the context of a negative first impression will lead to a failure to reverse the initial hypothesis, even in the presence of strong disconfirmatory, in this case, positive evidence” (p.115).

### 2.3 Mixing the Message

#### 2.3.1 The impact of Advertising

There is more than one way, however, to make a bad impression on a website. The visual content of a website is part of the aesthetic impression, and not all the content shown necessarily is relevant to the users expectation of what the website should be showing. The clearest and most common example of this is online advertising.

Advertising crosses the line between the aesthetic and symbolism aspects of user affect described by Rafaeli and Vilnai-Yavetz (2004). Not only does advertising have an aesthetic component, often itself at odds with the visual design of the rest of the interface due to intrusive animation (Yoo & K. Kim, 2005), but advertising may well incorporate messages and meanings that conflict, distract (Becker-Olsen, 2003) or are not relevant to the meaning of the main content (E. J. Newman, Jr, & Sprott, 2004), (Yus, 2005).
This tension has lead to a phenomenon known as Banner Blindness first discussed by Benway (1998) where users tend to simply not recall the advertising to which they are exposed online, although it has become clear that there is still a cognitive load placed on the user, reducing visual search (Burke et al., 2004) and scanning efficiency (Dahlén, Rasch, & Rosengren, 2003), slowing the speed with which users can interact with a website by interfering with the users experience of seamless flow (Sicilia & Ruiz, 2007) as well as adding to the overall download time of a web page.

Li, Edwards and Lee (2002) claim that online advertising is more distracting than in traditional passive media such as Television, because:

Consumers are often goal directed and may believe the ads are even more intrusive than when they are viewed in other media. They are likely to have negative attitudes toward ads or avoid them altogether to the degree that they feel the ads are unwanted. (p.37)

They go on to define advertising online as *noise in the environment* and illustrate how since users view ads as noise to be filtered out, the mere presence of advertisements within a web page affects the viewer's ability to reach their goals regardless of the nature of the advertisements, and “it is this placement that is likely to affect the degree to which the viewer perceives ads as intrusive” (H. Li et al., 2002, p. 39).

### 2.3.2 Advert Avoidance

Cho and Cheon (2004) found further evidence that advertising is generally avoided by users because of a perception of noise or clutter, but they also found evidence for the role of prior negative experience.

They confirm earlier studies, finding that the most compelling reason why users avoided ads and developed negative attitudes towards websites was that they felt there was too much “perceived” ad clutter, and that internet ads are a “significant source of noise or annoyance, hindering consumer efforts to browse Web content, they can disrupt consumer web page viewing, distract viewers from the web page's editorial integrity, and intrude on their search for desired information” (Cho & Cheon, 2004, p. 90).

That perception of “clutter” included any form of ad, banner ads, popups, massed groups of text links and advertorials. They concluded that “anything that impairs efficient interactivities between consumers and advertisers, such as placement, timing, and size of ads, can affect perception and be viewed as clutter” (Cho & Cheon, 2004, p. 93).
Another significant factor they proposed was prior negative experience with ads, which Cho and Chen suggested was “indicated by dissatisfaction and perceived lack of utility and incentive for clicking on those ads” (Cho & Cheon, 2004, p. 91).

Kelly, Kerr and Drennan (2008) built on this study with specific focus on social networking, and found that in that environment, word of mouth was a factor, as “most of the participants had not personally had a negative experience but knew someone who had” (p.4), and that second-hand knowledge proved enough to form negative impressions. Also, a degree of fundamental skepticism was observed, as “the participants in the study were distrustful of the information and motivations behind the advertising online” (p. 4).

Although there might have been several triggers for users avoiding ads, Cho and Cheon (2004) determined that there were three main behaviours and attitudes that resulted:

**Cognitive Avoidance** is determined by a consumers beliefs and thoughts about an advertisement. They suggest that this creates an intentional and considered response, and that “the more negative beliefs associated with Internet ads, the more unfavourable the overall cognitive component is presumed to be” (Cho & Cheon, 2004, p. 91), leading to intentional and conscious ignoring of ads.

**Affective Avoidance** is derived from a consumers feelings and emotional responses to an ad. Cho and Cheon suggest that consumers that intensely dislike ads are likely to form negative attitudes towards them, “and avoid the source of their displeasure” (Cho & Cheon, 2004, p. 91).

**Behavioural Avoidance** is defined by Cho and Cheon as the sum of the actions users engage in to get away from or skip over advertising, including “scrolling down Web pages to avoid banner ads, purging pop-up ads, clicking away from ad pages containing banners, and so forth” (Cho & Cheon, 2004, p. 91).

### 2.3.3 Irrelevant Added Content

This tendency toward ad avoidance creates a tension in the mingling of content and advertising on websites that is not always successfully resolved. Google, for example, has chosen to address this issue by a combination of subtlety and separation, as Spurgeon (2006) notes, “advertiser entries are incorporated into the overall flow of Google search results but they are also clearly distinguished from unpaid search results as ‘sponsored links’ that run down the right hand side of the search results page” (p. 6).
This conciseness and simplicity of presentation of both the advertising and search results “encourages end-user perceptions of Google as an accurate, authoritative, timely and comprehensive information source” (Spurgeon, 2006, p. 6), but in this Google is the exception that proves the rule, as their results page is designed with the sole purpose of presenting what is effectively a linear list of two streams of textual data; relevant results and relevant advertisements. It is a model of functional simplicity with minimal graphics that has changed little over the years. Most advertising and website content can not and does not knit together so easily.

The idea of advertisements being irrelevant or conflicting with the meaning of a website is developed by Newman et al. (2004), who draw on cognitive consistency theory to predict that “differences in affect and congruity of association between a banner advertisement and a Web site should have a strong influence on consumer attitudes toward a Web site” (p.274).

They give examples of congruous and incongruous advertising, e.g ‘luggage and travel’ or ‘shopping and credit’ versus say ‘travel and groceries’ and then go on to suggest that:

Consumers will experience tension caused by the incongruity of the two brands and will be motivated to find a way to either reduce the tension through a change of affect toward one of the two brands or by completely ignoring the message. The presence of a banner that the consumer perceives to be out of place with the Web page, therefore, may have a negative effect on the perception of the other brand. (E. J. Newman et al., 2004)

This negative impact from low relevance of advertising is potentially compounded by subjective social factors like unfavourable association with the content, themes or subtext of the advertisement. This highly subjective topic of advertising reputability is addressed by Fogg et al (2001), who pre-tested advertisements to select “one with high perceived reputability (an ad for Lexus cars) and one with low perceived reputability (an ad for gambling)”. Their study went on to demonstrate that “the low-reputability banner ad reduced perceived credibility of Web content significantly more than did the high-credibility ad” (p.2).

McCoy, Everard and Loiacono (2008) also explored this issue of advertising and credibility by comparing the affect of advertising on familiar and unfamiliar websites and determined that when ads were introduced, there was a significant difference in the perceived quality between the familiar and unfamiliar website. They found that “this demonstrates that the additional negative influence of ads on an unfamiliar site is more detrimental than their influence on familiar sites” (p.449).
It is clear that both website design and advertising have the power to communicate a variety of messages to the end user that may or may not match the designer's intent or fit the business plan of the website owner, and that this happens in ways that are largely subjective, poorly understood and difficult to measure exactly.

### 2.4 The Perception of Credibility

#### 2.4.1 Credibility and Trust

One of the strongest affective responses users have to websites regards their credibility and trustworthiness. Numerous studies (J. Kim & Moon, 1998), (Karvonen, 2000), (McKnight, Choudhury, & Kacmar, 2002), (Riegelsberger et al., 2004) have linked user affect with their assessment of website credibility. However, finding a satisfactory definition for concepts like Credibility and Trust in the first place has proven elusive.

There is a wealth of competing definitions, frameworks, models and three-letter acronyms relating to trust in the literature, with (Egger, 2000), (Gefen, Karahanna, & Straub, 2003), (Corritore, Kracher, & Wiedenbeck, 2003), (Chopra & Wallace, 2003), (Riegelsberger et al., 2004), (Riegelsberger, Sasse, & McCarthy, 2005), (Wang & Emurian, 2005) being perhaps some of the most cited examples.

Wang & Emurian (2005) suggest that there is some semantic confusion around the terms Credibility and Trust, in the sense that “trust is an abstract concept and is often used interchangeably with related concepts such as credibility, reliability, or confidence” (p.108). According to Everard & Galletta (2006) “trust is an attribute of an observer (to have trust), whereas credibility is an attribute of another person or an object of interest (to be credible)” (p.61).

This abstract concept of credibility can itself be usefully subdivided into more measurable component parts. Corritore et al (2003) identify four such dimensions of credibility: honesty, expertise, predictability and reputation.

They argue that credibility is “characterised by well-intentioned, truthful and unbiased actions” synonymous with honesty, and that it reflects expertise “typified by knowledge, experience and competence”. They also see the reputation of a website as capturing “the quality of past performance” and show that when “an object of trust acts in a predictable manner, credibility is lent to it” (Corritore et al. 2003, p.750).
2.4.2 Aesthetics and Credibility

Despite the debate about the exact definition of credibility itself, what has become clear is that the strongest single factor impacting user attitudes towards credibility of the website itself is the visual design of that website.

According to results of a large scale survey by Fogg et al (2003), user judgments on website credibility are 75% derived from a website’s overall aesthetics. They found almost 50% of comments about web credibility referred to some aspect of the design features of the websites, a figure that on first examination seems inexplicable. As they themselves put it, “one might ask, are people really so influenced by design look and not by more substantial issues? The answer appears to be yes” (Fogg et al., 2003, p. 12).

Others have echoed these sentiments. Karvonen (2000) wrote: “what may be surprising in this world of ours that so admires the assumed rationality of the homo sapiens, is that such an issue as trusting a service to be reliable and secure is, at least partially, based on something as irrational as aesthetics seems to be. If a Web site strikes me beautiful, I will gladly give away my credit card number - is that how it goes?” (p.87).

Other studies (McKnight et al., 2002), (Sillence & Briggs, 2007) have observed similar results, often noting a tendency for impressions to be made very quickly then being difficult to reverse, especially once a negative impression had been formed. According to Sillence & Briggs (2007) “poor interface design was particularly associated with rapid rejection and mistrust of a website. In cases where the participants did not like some aspect of the design the site was often not explored further than the homepage and was not considered suitable for revisiting at a later date” (p.31).

2.4.3 Credibility and Online Personas

This observed trend to use superficial aesthetic impressions as the foundation of sweeping social judgements becomes even more intriguing when real individuals, rather than abstract entities like a “website”, are at the focus of the attitudes formed. And the emergence of social media and social networks has thrust this issue into a wide range of online locations. Whether by intent or design, many online businesses, including some directory websites, create a limited and somewhat adhoc social network as a side effect of the main economic purpose of the website.

In many ways, these modern virtual replacements for the traditional main street or village market square have opened the door for wide scale social interaction between a greatly
expanded pool of participants, including sellers, buyers and bystanders, with communication freely persisting from the past to the future in a way that verbal conversations cannot.

As a vehicle of communication, the internet has given renewed emphasis to the power of word of mouth. According to Dellarocas (2003)

Not only can organisations reach audiences of unprecedented scale at a low cost, but also, for the first time in human history, individuals can make their personal thoughts, reactions, and opinions easily accessible to the global community of Internet users. Word of mouth, one of the most ancient mechanisms in the history of human society, is being given new significance. (p.1407)

Although the desire for self expression and recognition within wider communities has always existed outside the online domain, certain core features of online interaction and modern society in general lend new weight to these practices. Pescosolido & Rubin (2000) note that modern social networks resemble overlapping circles rather than the concentric, singular, community focus of the pre-modern era, with the degree of overlap and intersection in circles defining the individual socially. Social networks like Facebook and LinkedIn offer unique opportunities for self expression and relationship across social and physical boundaries, in effect, they provide tools for the feeding, care and maintenance of online persona’s (Gross, Acquisti, & Heinz III, 2005) that can be crafted to suit the context and culture of any given social network. These personas are often widely distributed across many sites, and can be inconsistent and contradictory. (Donath, 1999).

Importantly, it becomes very easy for individuals to have more than one online persona, each behaving in a manner appropriate to the particular social network. Negative and antisocial forms of interaction are effectively encouraged by the potential anonymity of online transactions on social network websites leading to participants showing a lack of responsibility for actions and consequences that would not emerge in normal face to face exchange (Alonzo & Aiken, 2004).

The diverse opinions, statements and potentially contradictory actions of an individual that form their various online persona’s usually remain visible to the wider web of users for extended periods if not indefinitely, and studies have shown that these online personas can leak into the real world lives of the participants (Gross, Acquisti, & Heinz III, 2005).

The very capabilities of the internet that enable diverse self expression online also assist others in examining the content of that self expression and drawing conclusions that may affect the offline life of the individual. Gross & Aquisti (2005) demonstrate that employers have begun using
the searchability and persistence of information on the internet to screen potential employees, and it has become increasingly common place to ‘google’ a potential partner or lover to investigate their suitability (Joinson, 2008).

In the case of online transactions, the potential for fraudulent, malicious or deceptive actions to be taken by either party to the detriment of the other or the wider community is an ongoing challenge that technology alone seems unlikely to resolve, and occasionally greatly worsens. Online shopping has been found to carry more risk than off-line shopping partly “because it relies on a complex sociotechnical system that stretches interaction over space and time” (Riegelsberger et al., 2004, p. 3).

This concept of the web as a sociotechnical system is an attempt to encapsulate and define the complex interplay between the technology of the web and the body of human communication and interaction it carries, and the changes and feedback this produces over time in both technology and the society in which the systems are embedded (Lyytinen & M. Newman, 2008).

According to Whitworth (2009), sociotechnical systems (STS) have arisen as a new communal system of exchange and interaction between individuals mediated by technology. This communal system is built upon the personal experience level of Human Computer Interaction, which is itself a refinement of earlier understandings of computing as mere combinations of physical hardware and information in the form of software.

These communal STS systems must ideally contend with the full range of human behaviour and often “face problems like mistrust, unfairness and injustice.” (Whitworth, 2009). The extent to which these systems fail to mediate the needs or expectations of interaction between people transparently or accurately is defined by Ackerman (2000) as the social-technical gap.

Such a social-technical gap exists with regard to many online businesses, as the interplay of the technology, user experience, and social context of online transactions may well have a broad
impact on the perceptions formed by individuals about others on both sides of the transaction, particularly with regard to their credibility and trustworthiness.

2.4.4 Reputation Systems

One common mechanism that has evolved to bridge this social-technical gap is the Reputation System. With the growth of social networking, enhancing or confirming reputation has become an increasingly conspicuous component of website strategy. Reputation Systems like reviews, ratings or user comments have appeared on many websites, providing more authoritative cues to credibility than can be conveyed by mere surface aesthetics.

Just as with human reputation, website reputation is a product of past actions being made visible to others. Toms and Taves (2004) suggest that “a websites reputation develops through positive previous exposure, through third party assessments such as rating services that have emerged on the Web, or indirectly through the linking of websites” (p.295). Unlike the direct visceral impressions formed from the aesthetics of the website itself, reputation systems permit the interactions, experiences and attitudes of previous users to influence new user’s responses and behaviour.

This persistence of past interactions into the current content of a website is a particular feature of many online businesses, ebay.com being perhaps the most famous advocate of a user review system as a fundamental part of its business model, but the potential for misinformation or incorrect communication is far greater in the anonymous medium of the internet than in face to face communication, due to “the volatile nature of online identities and the almost complete absence of contextual cues that would facilitate the interpretation of what is, essentially, subjective information” (Dellarocas, 2003, p.1410).

2.4.5 Warranting

In this environment of uncertain identity and intention, users and websites tend to develop conscious and unconscious strategies to better assess the reputation of a service or persona seen online. Some news aggregation services like Reddit.com and Slashdot.org employ variations on the concept of “karma” to help regulate the excesses made possible by anonymity, whereby the other users of the service are able to rate and rank the comments made by individuals, building a feedback system of reputation that is associated with an individuals persona across all contributions they ever make.
A more authoritative measure can be found in the concept of Warrants, which, as discussed by Warkentin, Woodworth, Hancock and Cormier (2010) are “connections between our real world identity and our online identity” (p.9).

They suggest that the usage of real names, photographs and real world acquaintances on websites connect personas encountered online to information grounded in real-world identities, and thereby “reduce the frequency of lies and constrain how serious they are” (Warkentin et al, 2010, p.9).

Amazon.com uses a policy of verified “Real Names™” as a means of adding veracity and status to the user generated comments that constitute a large portion of their web content. But where direct verification by the website itself is impractical, the task falls to the users discretion, which according to Warkentin et al (2010), tends to lead users to “trust other-generated remarks over self-generated remarks” (p.9).

Walther, Van Der Heide, Hamel and Shulman (2009) suggest that content that has been created or controlled by an individual about themselves provides “less warranting value than an institutionally based Web page that appears to be constructed by a webmaster or other third party” (p.232). Additionally, they found that users pay particular attention to negative information in impression forming processes, citing a number of earlier studies implying that negative information is more uniquely informative because it is more likely to reveal a person’s attributes whereas positive statements often reflect situational or societal standards (Walther et al 2009).

2.4.6 Symptoms of Trustworthiness

Where explicit external confirmations are unavailable to users, there is a tendency to sift secondary information for clues. This process is described by Riegelsberger et al (2005) as seeking Symptoms of Trustworthiness. These symptoms “are not specifically created to signal trust-warranting properties; rather, they are given off as a by-product of trustworthy actions” (p.392).

They give common examples of symptoms of trustworthiness, such as the presence of a large number of customer reviews for a product on an e-commerce site being a symptom of a large customer base. They also imply that “professionalism in site design and concern for usability” are symptoms of a “professionally run, and hence trustworthy company” (Riegelsberger et al., 2005, p.392).
These symptoms are a side effect of the website or organisations methods and functioning rather than a result of specific efforts. According to Riegelsberger et al (2005), this process yields benefits at no cost to the website, as untrustworthy competitors would need to mimic these same aspects to project the same impression.

The opposite effect also applies. Riegelsberger et al (2005) maintain that the “absence of symptoms of untrustworthiness (e.g. nervousness, scruffy looks, a poorly kept shop or an ineptly designed web interface)” (p.392), is a trust inducing sign that users depend on in their judgements of website trustworthiness.

The lack of clear signals seems to be no barrier to the formation of opinions in website users, there is evidence to suggest that users tend to arrive at some form of judgement about websites even where there is a shortage of information. Ellison et al (2006) give examples of small cues such as a user’s email address, the links on a person’s homepage or even the timing of email messages being enough for users to reach a decision about others.

This is reinforced by Braddy et al (2008) who draw on Spenses’ (1973) seminal work on Signalling Theory, to suggest that “when an individual does not have complete data, or is uncertain of the position he or she should take on a matter, he or she typically draws inferences based on cues from available information” (Braddy et al., 2008, p. 2993). According to Braddy et al (2008) available variables that do not seem to have any strong direct connection to the matter at hand can become significant cues that influence individuals attitudes where there is a lack of more definitive information.

2.4.7 The Anthropomorphic Web

An explanation for the degree to which signals as subtle as aesthetic appearance influence credibility both positively and negatively lies in the realm of human relationships. Tseng and Fogg (1999) in their seminal study of Credibility and Computing technology, refer to the concept of “Surface credibility”, which defines how much an observer believes someone or something based simply on inspection. They draw a parallel between human relationships and human computer interactions, declaring that “in the world of human relationships, we make credibility judgments of this type nearly automatically. The way people dress or the language they use immediately influences our perception of their credibility. The same holds true for computer systems and applications” (S. Tseng & Fogg, 1999, p. 42).

Chopra and Wallace (2003) explored this tendency to anthropomorphise computer interactions and found that it was “not necessary that the technology assume a sophisticated, human-like
persona” (p.5). They indicated that the social interaction was directed at the technology itself, rather than the person behind the technology.

Norman (2005) also observes this tendency, stating that “if something interacts with us, we interpret that interaction; the more responsive it is to us through its bodily actions, its language, its taking of turns, and its general responsiveness, the more we treat it like a social actor. This list applies to everything, human or animal, animate or non-animate” (p.137).

Seeing that aesthetics and other apparently minor cues can cause users to gauge the credibility of websites as they would a human social actor raises the issue of what it might mean to be associated with these websites. How much of this user affect is carried past the website to others that might be seen as part of their social grouping, the individuals, services and products represented on them?

The nature of such group context effects was explored in an experiment with Facebook Profile images by Walther et al (2008) which revealed that the “characteristics of other people who appear together with a target affect the perceived attractiveness of the target” (p.35).

When a person was seen to be associated in a group with more attractive members, the perception of their own attractiveness was enhanced. Equally, attractive people associated with unattractive members of a group were found to have their perceived attractiveness decreased. This process is also mentioned by Hassenzahl (2004), who identifies it as an example of an Assimilation effect, whereby “individuals in company of a beautiful friend are also perceived more favourably compared to being unaccompanied” (p.341).

The concept of trust transference may have some bearing on this issue. Stewart (2003) suggests that transference of trust can occur when an observer groups individual entities together either as a result of their proximity or because of their interactions with each other, and “that if one group member is believed to be trustworthy, when a new group member is encountered, that group member may be assumed to be trustworthy” (Stewart, 2003, p.7).

Stewart also argues that “during initial encounters, beliefs about a target (i.e trustworthiness) and beliefs about the situation in which the target is encountered (i.e the institutional factors in force) are yet to be differentiated” which may cause “a positive relationship between the two” (p.8). In the case of trust in a website, the website potentially fulfils the role of either a group or an institution, it becomes the larger context where the target, (i.e the service, product or persona) is considered and assessed.

From a social standpoint, these associative effects are perhaps not unexpected, just as being well dressed, well behaved or hanging out with bad company has social implications, these online
equivalents might easily be imagined to lead to similar conclusions in the minds of others. And since attractiveness and credibility have been already shown to be related in users affective responses, the possibility emerges that the same sorts of effects might be seen with more abstract entities like services and products depicted on websites.

### 2.5 A Multi Sided Exchange

The impact of aesthetics and advertising on credibility of websites have been reasonably well explored within the context of typical e-commerce or promotional websites as discussed above. In these environments the effects are confined to the direct relationship between a potential customer and whatever business the website represents. Both parties between them have responsibility for the entire transaction and all it communicates, for better or worse. They share the *system image* they have created, as Norman (2005) would have it.

The situation becomes somewhat more complex when the website is only an intermediary or a platform, the real transactions are taking place offline, and the potential communication involves two or more parties with very different objectives: *users, businesses* and *advertisers*, who can only perceive and assess each others behaviour, reputation and intentions through an interface and context that is beyond their control.

#### 2.5.1 The Two Sided Market Model

The overwhelming majority of studies into user affect and aesthetics have been made using examples that mirror the dynamics of normal e-commerce websites, with regard to the underlying business model. Most of these examples employ a traditional single sided business model, “Business to Consumer”, often shortened to B2C, where the consumer is using a website to survey or purchase the goods and services on offer from the merchant. This online shopfront may either replace or supplement a “brick and mortar” store (Mahadevan, 2000)

Most e-commerce websites use this model, where the only transactions supported by the website itself are those between the consumer and the merchant whom the website represents (Mahadevan, 2000). Other transactions, such as the logistics and supply that necessarily occur between merchant and wholesaler are executed offline, as it were.

In this model the merchant is effectively presented as the source of these goods and services from the point of view of the user of the website, effectively “disintermediating” the true transaction by removing redundant elements of the supply and distribution chain (Scott, 2003), (J. F. Mills & Camek, 2004).
But many more recent “Web 2.0” businesses, and some older online formats like “Business to Business” portals (B2B), auction websites, travel and business directories, exhibit a two sided market model based around a somewhat opposite “intermediating” effect.

The nature of this two sided market was made clear in Rochet and Tirole’s (2003) seminal work on platform competition in two sided markets. They suggest that most marketplaces with network externalities, where both the customers and suppliers can use the business as an intermediary to operate an exchange tend towards this model:

Buyers of video game consoles want games to play on; game developers pick platforms that are or will be popular among gamers. Cardholders value credit or debit card only to the extent that these are accepted by the merchants they patronise; affiliated merchants benefit from a widespread diffusion of cards among consumers. (Rochet & Tirole, 2003, p.1)

Success in this market format requires the intermediary or platform business to take care to “get both sides of the market on board” (Rochet & Tirole, 2003, p.1). Rochet and Tirole claim this involves careful consideration of pricing structures rather than simple price level for the service offered. Which side of the market to charge for the services provided becomes a complex issue, as “platforms often treat one side as a profit center and the other as a loss leader, or, at best, as financially neutral” (Rochet & Tirole, 2003, p.2).

A major factor in this assessment of pricing structure is a clear understanding of the specific network effects in operation for this marketplace.

Network effects rose to prominence with Katz and Shapiro’s (1985) ongoing studies into the mechanisms by which “the utility that a user derives from consumption of the good increases with the number of other agents consuming the good” (p.1). Although somewhat popularised...
and oversimplified into buzz word status over the last twenty five years, the impact of indirect network effects is vividly portrayed in two sided markets.

Jullien (2005) illustrates how these effects produce a classic example of the chicken and egg problem for intermediary businesses, as “a customer on one side of the market will be willing to participate to the platform activity only if he expects a sufficient participation from the other side” (p. 3). Parker and Van Alstyne (2005) found that intermediaries who can play across both sides of the market can often structure their pricing in ways businesses that only serve either market separately can not, giving them a competitive edge and helping to drive adoption of their platform for both sides of the market, as “growth on one side of a matched market then induces growth on the other, creating exploitable surplus” (p.1497).

2.5.2 Online Business Directories

One common example of such two sided markets on the web are Business Directories. Even traditional printed Yellow Pages have been noted as clear examples of two sided markets and network effects (Rysman, 2004), and many of these same services have developed online versions. But other niche forms of business directories, such as travel, automotive and health related directories have also flourished on the web.

All these directories share common features that make them distinct from portals, auction websites or e-commerce sites, many of which appear functionally similar. Directories tend to function by aggregating a number of competing commodity services by location and type, and in many ways, directories both online and in print resemble search engines with a limited scope, and can arguably be considered early examples of search technology (Spurgeon, 2006).

Directories, according to Lohse and Rosen (2001) are also examples of “Directional Media”. Unlike traditional media which tries through advertising to build an image or brand for a product or service that may lead to an eventual sale at some later time, directional media’s purpose “is instead to develop immediate action either in the form of a telephone call or a store visit. When consumers come to the Yellow Pages, they are ready to make a purchase and are looking at the advertising specifically to select a supplier” (Maher 1988 as cited by Lohse & Rosen, 2001, p.73).

Importantly, unlike auction or e-commerce websites, directories are mediating a financial exchange or transaction that is not completed via the actual website, but offline or on another website altogether. As a result they cannot employ commissions or directly profit from the transactions.
Directories tend to favour keeping the consumer side of the market as a loss leader as anything which might impede consumer usage breaks the positive network growth effect (Rysman, 2004), so the alternative for most directories both online and in print is to charge fees to the business services side of the marketplace, who’s willingness to pay is directly dependant on how many consumers are using the directory (Rysman, 2004).

The economies of scale and eventual commoditisation of internet hosting technology has significantly lowered the cost of entry for intermediary businesses compared to their brick and mortar equivalents (Beuscart & Mellet, 2008). Additionally, the low operational costs of information distribution compared to physical product distribution enables online businesses to scale up easily and cheaply by subsidising “an arbitrarily large market based solely on fixed initial costs. Providing each new sale or service then costs the clever product designer nearly nothing in incremental costs” (Parker & Van Alstyne, 2005, p. 1497).

A key effect of these low barriers to entry and distribution online is an increase in the number of competing online directory services entering the market, which directly benefits business services by driving down prices and benefits consumers because they “view directories as differentiated products and so more directories attract more consumers” (Rysman, 2004, p. 35)

This competition between intermediaries benefits both sides of the marketplace but reduces the profitability of individual directories themselves (Rysman, 2004).

To compound the downward price pressure on directory platforms online, consumers also benefit from the low cost of discovery and low barriers to change that the web provides by exploring multiple options for purchase. This lowered “opportunity cost” (Jullien, 2005, p. 4) for users means that “a fraction of end users on one or the two sides connect to several platforms. Using the Internet terminology, we will say that they multihome” (Rochet & Tirole, 2003, p. 5).

So, with little to enforce consumer loyalty, no means of billing consumers without impacting network effects and a low cost environment where multiple competing directories drive down the profits from business services paying for listings, Online Business Directories often find themselves turning to more indirect sources of income that must be obtained outside the two sided market they are servicing.

### 2.5.3 A Love - Hate relationship

One obvious and very common choice of income for websites that struggle to monetise either side of the marketplace they are serving is Banner and Adword Advertising. Anderson (2009) rather wryly observed that “the standard business model for Web companies that don't actually have a business model is advertising” (p. 3).
Interestingly, for online directories, mixing paid advertising like banners and adwords with directory listings, which are also a form of advertisement, means integrating two potentially competing two sided markets into one product, effectively creating a **Multi Sided Online Market**, as depicted in Illustration 4.

Consumers of advertising are often split into two camps, Ad-haters and Ad-lovers (S. P. Anderson & Gabszewicz, 2006). Ad-haters are exemplified by the typical consumer of broadcast media, magazines or online news and information portals. These consumers are often presented with so many attention consuming advertising messages that they “have to be paid (or forced) to consume it. Most TV or magazine ads are consumed because consumers are paid with the programs they go with” (Beuscart & Mellet, 2008, p. 7).

Ad-lovers on the other hand “may find a positive net benefit from ads” (S. P. Anderson & Gabszewicz, 2006, p.5) as is the case in directory listings where the consumer is actively seeking advertising to address a specific perceived need. This sort of advertising will therefore attract a certain subset of consumers. However, this available pool of ‘eyeballs’ is itself an attractive and marketable commodity to other advertisers wanting to get their message across, “regardless of whether the readers or viewers are attracted to the messages per se” (S. P. Anderson & Gabszewicz, 2006, p. 5).

So online directories that choose to also display other forms of advertising may create conflicted consumers. Those ad-loving consumers who are willingly using one form of advertising in a directory in pursuit of specific information may find themselves also confronted with less
relevant or possibly annoying advertising that consumes their attention and provides no useful information (Beuscart & Mellet, 2008) pushing them toward the role of ad-haters.

The relationship between the perceived quality of the website, the value of its content and the quality and relevance of the advertising displayed is also a factor that online directories must balance carefully. For many advertising distribution networks, “associating brands and advertisements with "premium contents" is part of their definition of the quality of the service” (Beuscart & Mellet, 2008, p. 12).

These networks will only run advertisements in carefully selected environments to protect the brand image of their clients. “Just as luxury goods only advertise in high-standard magazines, these ad-networks guarantee that brands will not be associated with ugly, low-standard or questionable contents” (Beuscart & Mellet, 2008, p. 12).

Some free online directories, being highly competitive and needing to maximise content to maximise viewers, often run numerous unpaid listings; effectively free advertisements for services that may or may not be maintained by the businesses involved. This potentially uncurated content has a tendency to diminish the perceived relevance, accuracy and credibility of the listing information, which in the eyes of the advertisers, diminishes the “premium” value of the website (Beuscart & Mellet, 2008, p. 13) and increases the likelihood of low quality or irrelevant advertising being deployed instead.

Adword type advertising, where the advertisements shown are automatically selected, or targeted based on search terms found in the content can also lead to a downward pressure on quality as according to Beuscart and Mellet (2008) the matching system will “display the related ads without making any difference between high and low standard contents” (p.12) and although the “advertiser can choose the keywords he wants to associate his ad with, he cannot discriminate between the sites that match these keywords” (Beuscart & Mellet, 2008, p. 12)

Revenue from advertising is usually determined by page views and is maximised by exposing the most content to the most users rather than attempting to provide the best content (J. Hamilton, 2004). So directories depending on this income stream are not motivated to provide better tools for comparing different listings like professional reviews or other curated content, instead they seek to expose as many individual listings, wrapped with as many ads as possible to the largest audience they can muster for the longest time possible (S. P. Anderson & Gabszewicz, 2006).

Directories taking this approach have to walk a fine balance to satisfy the needs of the ad-loving consumer in search of information without turning him into an ad-hater by the sheer volume of
secondary advertising that comes along with the useful information. The only meaningful feedback on how well an ad-supported directory can achieve this balance and prosper is in their successfulness at building a loyal critical mass of returning users (Beuscart & Mellet, 2008).

2.6 A Multi Sided Network of Affect

All the elements discussed in this literature review can now be collected into a broad sociotechnical model of the mechanisms by which affect and credibility may be related and transmitted in a Multi Sided Online Market. The sum of these potential transactions and permutations have been formulated by this researcher into a proposed new model, which we have framed as the Multi Sided Network of Affect. A diagram showing the potential relationships and stimuli in this model is seen below in Illustration 5.

![Illustration 5: Multi Sided Network of Affect](image)

This Multi Sided Network of Affect, with users on one side and services and advertisers on the other, is united in its transactions and communications through the intermediary platform website, e.g a business directory.
All impressions and affective transactions between participants pass though the filter of content, advertising, user reviews and aesthetic presentation that is the sum online presence of the directory. These collected affective impressions are examples of the merging of Rafaeli and Vilnai-Yavetz’s (2004) Aesthetic and Symbolic forms, or of Normans’ (2005) Visceral and Reflective Designs. Together, these factors combine to form the meaning and significance of the directory to all participants.

Regardless of whether the significance and style of the presentation conflicts with message of the core content, communication of meaning is inescapable, and can be expected to leave a unique affective impression that may differ for each participant on all sides of the market. Rafaeli and Vilnai-Yavetz (2004) make it clear that “the symbolism of an artefact regards the multiple associations the artefact can represent, not only the associations formally intended by an organisation. Emotions can and are likely to be evoked by any association, and emotional pitfalls lay in failures to recognise this” (p.104).

In this Multi Sided Network of Affect model, all the participants can observe or experience aspects of each others online persona through the collective meaning of the website, and the potential relationships can become complex. Users clearly can view and form impressions of the services on offer, this is a baseline effect of any directory, but services may experience reviews and feedback from users which gives insight in the other direction. Additionally, because all the information is visible to all participants and preserved over time, services can see how their competitors are perceived, and users can gain impressions of their peers on the directory through the content of their reviews.

Advertisers too, can assess the quality of the presentation, services and user participation to determine the type of advertising they deem suitable, indeed they can also assess what competing advertisements are being supported, and consider if they belong in association with that group.

Impressions of credibility and changes in emotional affect between participants may therefore be expected to potentially travel in all directions though the intermediary, and also to be occasionally mirrored back in the presence of reviews, feedback and direct comparisons between services and advertisers.

This imagined Online Business directory then, is no longer just a Phone Book. The whole network can in principle be considered something of a community that exists within an intermediary neighbourhood, where individuals may form impressions about others around them, and also about the environment in which they are set.

Literature Review
2.7 Bad Neighbourhoods on the Web

How may we then determine if it is a bad neighbourhood or a good one?

While it seems clear from the literature that the perceived credibility of a directory website itself will be impacted by its aesthetics and content relevance, it is an open question whether this affective response will be transferred between the users and services, products or personas depicted on the site.

It is this gap in knowledge that this research is intending to explore.

There are four possible vectors of affective impressions of credibility, just between users and businesses, that are potentially supported in the model described above:

- User's impression of Service credibility
- Service's impression of User's credibility
- Service's impression of competing Services credibility
- User's impression of fellow User's credibility

Examining them all in a quantitative fashion is beyond the scope and capability of this research, so this study will limit itself to examining the User to Service vector, and attempt to determine the extent to which the perceived credibility of a web business directory affects user perceptions of the credibility of the services represented on it.

The hypotheses being examined are:

- **H1.** Reduced website attractiveness negatively affects users perceptions of the credibility of the services represented by the directory.

- **H2.** The presence of website advertising of low relevance negatively affects users perceptions of the services represented by the directory.

- **H3.** The presence of website advertising of low relevance and low reputability negatively affects users perceptions of the services represented by the directory more than low relevance alone.

Additionally, emergent issues around the role of reputation, aesthetics, user affect and advertising on the users perceptions of all parties in this Multi Sided Network of Affect will be explored in a qualitative component.
3 Research Design and Methodology

3.1 Research Methodology Overview
To compare the effects of different kinds of website experience on credibility and address the hypotheses, a series of experiments will be outlined that will use an online research website and instrument.

3.1.1 Research Website and Instrument Design
A fictitious online business directory website will be employed to examine the effect of different kinds on website experience on the users perception of the credibility of services encountered on the website.

The independent variables for this research are:

IV1. Website Design Attractiveness
IV2. The presence of Advertising of low relevance

Each of these independent variables explores one of the hypotheses H1 and H2 described above. A 2 x 2 design will be employed, consisting of two variations each of two independent variables for a total of four versions.

Version 1 will be a “control” website, simulating a basic business directory website. It will display an 'Attractive' website design, and no advertising.

Three other versions of this website will be created, each of which will be identical to the control version from a content perspective, but modified to present a different combination of the two variables.

Version 2 will have an 'Unattractive' visual design. It will have no advertising.

Version 3 will have the “Attractive” design, but it will also contain advertisements of irrelevant and unrelated content and services.

Version 4 will combine both the 'Unattractive' design and the irrelevant advertisements.
Each service will be presented on its own webpage in a manner consistent with other example online directories, but that page will also contain an online survey form designed to assess the effect of the various versions on the perceived credibility of the service.

Details on the differences between the Attractive and Unattractive designs, the nature and origin of the advertising and the mechanics and details of the experimental instrument can be found below from section 3.2 onwards.

### 3.1.2 Pilot Study

A small questionnaire based pilot study was used to validate the main experimental assumptions about the relative attractiveness of two website designs, and the relative lack of appeal, irrelevance and disreputability of advertising used in the research instrument. The methodology and results of this pilot study can be found below in section 4.

### 3.1.3 2 x 2 Between Groups Study

A 2 x 2 between groups study was conducted using the research instrument described above, in which a baseline control version of an Online Directory with 16 services is compared to 3 versions that cover the other possible combinations of independent variables to see what effect these variables have on the dependant variable when each group of participants only experience a single version of the website.

Participants were randomly assigned to just one of the four variants for their entire session, and were asked to complete a brief online survey capturing their impressions about each page of the website they view. The methodology and results of this between groups study can be found below in section 5.

### 3.1.4 2 x 3 Within Subjects study

A 2 x 3 within subjects study was conducted using a modified version of the research instrument described above that permitted two versions of advertising to be presented, one for irrelevant advertising, and the other for irrelevant and disreputable advertising.

Each participant was randomly assigned to a counterbalanced combination of 6 services and experimental conditions as outlined below:
Attractive design + no advertisements
Attractive design + irrelevant advertisements
Attractive design + irrelevant and disreputable advertisements
Unattractive design + no advertisements
Unattractive design + irrelevant advertisements
Unattractive design + irrelevant and disreputable advertisements

This design was used to confirm the results of the first experiment with regard to the hypotheses H1 and H2 using the greater power of a within subjects design.

It also enabled the assessment of Hypothesis 3:

H3: The presence of website advertising of low relevance and low reputability negatively affects users perceptions of the services represented by the directory more than low relevance alone.

The methodology and results of this within subjects study can be found below in section 6.

3.1.5 Qualitative Analysis

The user comments in open ended questions in both studies were coded for thematic content and the results studied to determine the prevalence and relatedness of the emerging themes to aspects of interaction issues, aesthetics, user affect, content meaning, credibility and reputation, assessing for negative or positive projections. The results of this work can be seen in section 7.

3.1.6 Previous Research of this Nature

These basic research methods are similar to a number of prior studies, even in regard to the use of an online survey approach compared to face to face surveys (Coursaris et al., 2008) for research related to aesthetics and credibility. The use of both between groups (Coursaris et al., 2008), (Fogg et al., 2001) and within subject (Brady & Phillips, 2003), (Tuch et al., 2009) experiments was seen in many of the studies examined in the literature review.

The qualitative approach used is drawn from examples by Fogg et al (2003), Alsudani & Casey (2009), and Michailidou et al (2008).
3.2 Research Website Design

3.2.1 Gathering example designs

A total of 23 directory and e-commerce websites with diverse appearances and aesthetic appeal were examined to assess recurring themes and stylistic aspects that would be employed to create the Attractive and Unattractive Designs. A full list of source websites can be seen in Appendix 1.

These websites were largely retrieved via search engine results, although some of the New Zealand business directories were already familiar to the researcher. Where possible, business directory listings were preferred, but functionally similar e-commerce pages that exhibited interesting aesthetic qualities or weaknesses were also included.

From these, 12 websites that displayed identifiable characteristics in keeping with aesthetic aspects explored in the literature review were used as references for the study design as seen in Illustration 6 below.

![Illustration 6](image-url)
In order to avoid ongoing changes in the websites, full length screenshots of the page content were taken using Firefox 3.6 on Mac OS X and a firefox screenshot plugin called Screengrab! available from http://www.screengrab.org/

These screenshots were all taken in November 2010 with the browser window set to 1280 pixels wide and are included in Appendix 1.

### 3.2.2 Design Principles

The 12 websites were chosen for specific positive and negative features regarding the following aspects of their designs and their impact on the perceived complexity and appeal of the web page as discussed in the literature review.

**Colour:** The specific wavelengths of light reflected by objects and backgrounds, and their relationships in Colour Theory.

**Contrast:** The differences and relationships in intensity of colours, from dark to light, and from grey to saturated.

**Shape:** The discrete volume that a visual object occupies, and the nature of the edges, the distinction between an object and its background.
Texture: The discourse of tones and transitions perceived to be within an object, the implied surfaces and dimensions of an object or background.

Positioning: The way objects are aligned, the apparent spaces created, orientations, rotations, the implied relationships between things across the entire visible frame.

Consistency: The ability to relate the meaning of multiple items by their similarities and differences in proximity, pattern, shape, texture, colour or contrast.

Flow: The movement of the eye over the total design, the priority and emphasis given to various items over others, the way items succeed or fail to work together in a design.

3.2.3 Applying the Design Principles

Colour

Two examples used for positive colour guidance were shopnewzealand.co.nz and finda.co.nz. These two sites employed two different implementations of Colour theory (Coursaris et al., 2008). Finda used Complimentary colours, blue and orange, whereas Shopnewzealand used Analogous colours, a deep blue and an aqua, both of medium intensity.

Additionally, these key colours used were derived from the brand/logo colours for each site. Rather than mixing the brand with colours that conflict, the colouration of the logo is carried across the design.

This colour usage was applied to the Attractive design, where similar green and blue shades of medium intensity were used in the logo design and then carried across the page. Orange was used as a highlight for review stars as seen in Illustration 7.
For the Unattractive design, *Looksie.com*, which was the inspiration for the domain name used in the research, gave a great example of a too-colourful logo, using a rainbow effect and a cartoon like eyeball metaphor. This use of the full spectrum of colours leaves no direction for the branding of the site, any and every colour is related, so the effect is to single out no particular theme.

The directory results page at *goldenlanterns.com* and *topmaq.co.nz* gave inspiration with their broad usage of intense red, deep blue and bright yellow. These complementary primary colours, when used in close juxtaposition, create a perceptual effect known as Vibrating Boundaries, which appears as an illusion of movement or flickering in the edges of text and graphics that are embedded in strongly contrasting surrounding colour backgrounds. This brandless approach to colour, use of strong primary tones and cartoony logo was used in the Unattractive design as seen in Illustration 8.

![Contrasting Primary Colours](image1)
![Full spectrum of colours](image2)
![Unattractive Design](image3)

**Illustration 8: Colour use in the Unattractive design**

**Contrast**

In *shopnewzealand.co.nz* and *finda.co.nz*, care has been taken to ensure that the contrast between the colours is controlled, by having intensely different colours appear widely scattered on a pale or white background to draw the eye to a few important features like titles or specific icons, rather than using them excessively in close combinations.

Additionally, neither of these sites use repeating decorative patterns or any background texture than would reduce contrast with the text. The background is either white or a very pale graduation, providing maximum legibility to the body text. This same approach is carried over to the Attractive design as seen in Illustration 9.
The homepage of Goldenlanterns.com, on the other hand, shows excessive contrast between the text and the background colour, rendering the text almost unreadable. Contrast can add to the overall visual complexity of a design by having multiple edges, lines, boxes and shapes in close proximity, effectively producing alternating striped or strobing areas of design that automatically draw attention (Treisman, 1985) away from the wider content. Examples of this can be seen in the navigation side bar of looksie.com, the alternating dark blue and white panels of goldenlanterns.com search results and to a lesser extent the white text and spacer lines in the tradeboss.com navigation bar. This alternating line and high contrast text effect is replicated in the navigation elements of the Unattractive design as seen in Illustration 10.

Shape

Shapes of many objects on the web tend towards rectilinear forms, a consequence of the underlying limitations of the html technology, an artefact of the screen shape, and a vestige of the cultural history of graphic design in print media with it’s underlying emphasis on lines and columns of type.

The typical exceptions to this rectilinear tendency are organic and naturalistic shapes found within images and large font letter forms.
For every object, there is also a potential space around it. This is what is often termed the whitespace or negative space between objects or collections of objects. These are essentially areas of the background that are visible between objects, or between rows or columns of text. The choices of shapes and their arrangements form an unconscious tension between background and foreground in the perception of the viewer. This balance is often termed the figure-ground separation (Marr, 1982), (Treisman, 2002).

The negative effect of whitespace imbalance is neatly and literally demonstrated by topmaq.co.nz, where their product pages often have irregular patches of white background appearing that are almost as large as the meaningful content area.

Even yellow.co.nz demonstrates a tendency towards poor use of available space and an excess of background, partly as a result of not addressing browser screen size differences in a flexible way. Their attempt to reuse the brand values of curved shapes from the logo has resulted in gaps, unusual orientations and alignments in the navigation bar that makes the functionality of the navigation become divided and incoherent.

The Unattractive design attempts to introduce odd, random and unbalanced whitespace areas across the whole page as seen in Illustration 11.

![Illustration 11: Unbalanced shapes in the Unattractive design](image)

**Texture**

Background textures and pseudo-dimensional effects like glows and drop shadows are often employed to give an illusion of layering to the two dimensional presentation of a web page. Careful use of small amounts in a consistent fashion can help increase clarity by producing greater apparent separation between foreground objects and their background, and allowing multiple items to find their own level relative to one another.

A small amount of shadow effects, graduations and dimensional effects can be seen on the boxes and buttons of shopnewzealand.co.nz and finda.co.nz, and the result is a pleasing but unobtrusive means of drawing attention and adding perceptual space to objects. Similar effects were applied to similar elements in the Attractive design as seen in Illustration 12.
The use of strong background patterns and textures behind text by sites like countrycreekconnection.com and tamilbizcard.com diminishes the contrast and legibility of the foreground text, either by insufficient tonal range or adding an excessive number of hard shapes and edges that conflict with the letterforms of the text. This usage of background textures was much more prevalent in the early days of the web, but can still be found in places. A background texture has been added to the Unattractive design as seen in Illustration 13.

Positioning

Spaces or objects that appear similar but not exactly the same in close proximity or in a set appear as errors and draw attention unconsciously (Treisman, 1985). Shopnewzealand.co.nz and finda.co.nz are careful in their alignments and orientations of elements within blocks of content. Objects in close proximity share at least one and often two vertical or horizontal edges, or in some cases use a centreline. Items in a list will typically have the same spacing between elements and share an alignment on at least one edge. Use of sufficient spacing or margins helps allow the eye to concentrate on small collections of parts of the design as units rather than trying to comprehend the entire collection of items simultaneously, reducing the sense of clutter and improving the speed of scanning a block of content. (Doll, 1993).
Tradeboss.com and goldenlanterns.com on the other hand, show many examples of irregular spacing and alignment of objects in close proximity. Some text blocks are centred, others left aligned directly under each other, boxes and edges are at times close but not close enough to be unified, drawing attention to these aspects as if they were meaningful errors.

Another more subtle issue revolves around vertical and horizontal spacing of objects inside containers, a design aspect often termed as padding. Tradeboss.com and goldenlanterns.com show numerous locations where text in containing boxes lacks space, appearing to practically collide with the borders and edges of the containers, and in other places have widely divergent variations between the padding of multiple objects in the same container. These subtle near misses and close spaces are unconsciously but swiftly noticed (Treisman, 2002), causing the elements in question to draw more attention than intended.

Line spacing of text, or leading, has a well studied effect on legibility, too little or too much impedes reading and can affect the overall sense of clutter or clarity of a page (Rosenholtz, Y. Li, & Nakano, 2007) (Doll, 1993), and tradeboss.com and goldenlanterns.com again show this effect in a number of locations. Goldenlantern.com creates numerous illegible blocks of links, as legibility issues from inadequate line spacing are compounded by the presence of the underline on most conventional text links. Tradeboss.com has many examples of very densely packed body copy, with hardly any paragraphs to split the message into manageable chunks for easy reading. Spacing irregularities like these have been introduced into the design of the Unattractive website as seen in Illustration 14.

Positioning in the broader context, often called Composition, addresses the way blocks of objects within the page relate to each other and the total design. Groups or chunks of objects with spaces between them are able to be perceived as if they were fewer, larger single objects, reducing the apparent complexity of a design.

In many cases specific effort has to be applied to position objects in html, Goldenlanterns.com homepage particularly demonstrates the sorts of random groupings that can occur when design
is mainly left to the natural page flow in html browsers. These random spacings and alignments are interpreted as meaningful even when they are not or can mask meaningful arrangements that do exist by introducing clutter (Rosenholtz et al., 2007). Some of this randomness of composition has been introduced into the Unattractive design as seen in Illustration 15.

Illustration 15: Random positioning in the Unattractive Design

Consistency

Shopnewzealand.co.nz and finda.co.nz use fonts and colours consistently. They do not mix serif and san serif fonts at all, and certainly not at the same sizes. Bold and italic forms are used sparingly and consistently, and font sizes are kept uniform within a given block of content. The same sort of items used in the same sort of way receive the same font, styling, colouring and sizes. e.g text links in a block or items in a list a consistent.

Both these sites use san serif fonts (e.g Arial, Helvetica) at a moderate 11 - 13px size as body copy. Traditionally, serif fonts (e.g Times) were used in print for body copy as they provide higher legibility, but on screen the serifs add visual complexity at small sizes due to the resolution limitations of monitors, especially for older readers (Bernard, Chaparro, M. M. Mills, & Halcomb, 2003). So the resulting san-serif blocks of body copy are very legible. A consistent approach to colours and fonts types was used in the Attractive site as seen in Illustration 16.

Illustration 16: Consistent Fonts & Colours in the Attractive Design
Mixing serif and sans serif fonts, bold and italics and slightly different sizes of fonts in a given area, as illustrated in a number of places by tradeboss.com, reduces clarity by undermining the distinction between headings and body copy, or reducing the grouping effect that consistent lists produce. This sort of font variance was used liberally in the Unattractive design as seen in Illustration 17.

**Illustration 17: Inconsistent Fonts & Colours in the Unattractive Design**

**Flow**

Two aspects of flow need considering in web design. One aspect is the implied hierarchy or sequence of items that the viewer is expected to perceive. This aspect is common to all static print media, and involves the strengths and sizes of headings, subheadings, and blocks of colour. Having good distinction between headings and body copy or using lighter coloured regions to bring the objects forward and darker ones to push them back (Treisman, 2002) are design features used by shopnewzealand.co.nz and finda.co.nz and are reflected on the Attractive design.

Confused flow leaves a design with no clear overall hierarchy of purpose and order of reading, as seen in yellow.co.nz’s navigation bar where the swooping shapes lead the eye on a wandering journey through a complex white space with no clear destination. Goldenlantern.com also suffers from this problem, their homepage is a complex mass of text links and icons, with various alignments and spacings and no clear sense of priority or internal flow. Everything is competing violently with everything else equally, and no clear winner emerges. The flow of the Unattractive site was made deliberately confused by mixing heading and body font sizes and colours across the entire design.

The second aspect of flow, not encountered in print media, is the distracting effect of Animation. Animated items on web pages draw an excessive amount of attention because of their often unpredictable motion. These sort of distracting movements occur without any stimulus from the user, as opposed to useful interaction feedback animations like rollovers and progress bars that usually occur at or near the location the user is focused on.
Animation that occurs at random times and locations on the page draws attention away from the content the user was attempting to examine toward objects which often have no particular intrinsic meaning or message to offer. Moving items can become dominant as a web page is otherwise essentially static, so the addition of movement has high impact. Items like animated flags, logos, dancing cartoon characters and icons are common examples of this sort of distraction found on the internet.

Goldenlantern.com exhibited numerous independent animated icons cycling continuously on their homepage. None of these icons had any particular apparent reason to dominate the visual space, the choice of animations seemed arbitrary, and the number used meant that it was difficult to focus on any part of the page long enough to understand its purpose. To capture some of this effect without dominating the design utterly, a single animated icon of a waving New Zealand flag was added to the Unattractive design as seen in Illustration 18.

Illustration 18: Animation in the Unattractive Design

3.3 Research Website Content

3.3.1 Types of services listed

A fictitious website was designed simulating a New Zealand business directory website. In order to give a more realistic sense of a directory, four types of services were included:

- Accountant
- Plumber
- Restaurant
- Hairdresser

These four service types were chosen to be very mainstream, to ensure that the majority of participants would have some pre-existing familiarity with the nature and function of the services, despite their age or whether they were home or vehicle owners. The choice of
restaurant and hairdresser particularly was to ensure that the overwhelming majority would have had some direct experience of these types of services.

Accountant was added to allow a wider range of business service types to be depicted, where the majority of plumbers, restaurants and hairdressers in New Zealand might be typically owner operated or small business, having accounting firms gave the opportunity to add a more corporate aspect to the study.

3.3.2 Content of each listing

The services of each type were be presented in a consistent format; a fictional search results detail page with the following core data presented:

- Company Logo
- Company Name
- Service Overview
- Location Map

A proportion of the services additionally had:

- Example Images
- User reviews

These elements were chosen after study of commonly recurring items found on example detail pages in online New Zealand Business Directories. The directories examined were:

- www.finda.co.nz
- www.yellow.co.nz
- www.zipleaf.co.nz
- www.hotfrog.co.nz

Some of these websites also provided design inspiration as discussed earlier, but in this context it was the structural elements of the content that was being considered.

3.3.3 Gathering sample data for the listings

The textual, image and logo data, including descriptions for these 16 businesses was collected mainly from the Australian Yellow Pages (www.yellow.com.au) during November 2010.
Using the Australian directory was a means of finding company names, images, logos and details that would be unlikely to be duplicated in New Zealand but had enough cultural similarities to be very plausible. Phone numbers, urls and address details were fictionalised and a fake Google map was prepared for each.

A large body of data was collected, then some care taken to select a subset to represent a wide range of apparent business size and professionalism for the services. This sense of a range of service size and quality was partly a reflection of the richness and depth (or lack thereof) in the original content, but was also enhanced by the deliberate addition of supporting images and user reviews to some but not all services. The objective was to produce a distinct and somewhat random set of different perceived services, rather than some graded hierarchy from best to worst. Since the independent variables being assessed in this study were across the aesthetics of the presentation and the presence of advertising rather than variations in the listing content, the individual responses participants might have to individual services relative to others in the same user session was not significant.

The user review information was mostly gathered from www.fnda.co.nz, one of the few directories that support user reviews. The reviews were taken from completely different businesses and in some cases different service types altogether, and many were anonymised further as required.

Importantly, existing spelling mistakes, capitalisation and both negative and positive assessments of debatable merit were used largely unmodified, to try and more closely simulate a realistic set of user responses.

### 3.4 Adding Advertising

Two more versions of the website were created by adding advertising to the existing Attractive and Unattractive versions. This advertising was selected from actual advertisements found online during December 2010 and was reduced down to five sets of display advertisements, and five sets of adword groups.

#### 3.4.1 Gathering example advertisements

Examples of online advertisements are easy to find, and some of the websites used as example designs contributed to the search. Numerous sites were evaluated, but since the advertising seldom was directly associated with the sites, could be found in multiple places simultaneously and was usually rotated on each page view, there is no point in specifying the exact locations.
However, certain types of website contributed the best material. News Aggregators like Digg.com and Slashdot.org provided a number of ugly but essentially innocuous advertisements like teeth whitening and weight loss. Casino advertisements, “bait and switch” special offers and online dating advertisements were often found on second rank aggregators and entertainment sites like Break.com, Mgid.com and theChive.com. Also Google Image was used to provide a rich and context free source of animated and banner advertisements.

The same widespread web surfing approach was taken with gathering suitable adwords advertisements, which were common on many of the example directory websites, and they often demonstrated either a very oblique or nonexistent connection between the page content and the advertising themes and services.

3.4.2 Sizes and Types of Advertising

Common usage on the Internet has evolved numerous uniform “Industry Standards” for sizes and types of graphic advertisement that can be deployed to various websites automatically. These were formalised into the Universal Ad Package (UAP) by the IAB in 2002 and consist of four sizes:

![Illustration 19: Four common advertising sizes](image)
According to the 2009 review of the IAB - Ad Unit Guidelines, “the UAP has been widely adopted and the four sizes still account for the vast majority of impression weight in the marketplace”

For the purposes of this research, the Leaderboard and the Medium rectangle were selected as the sizes of display advertising that would be used on two versions of the website, the Attractive design with advertising and the Unattractive design with advertising.

A feature of this particular combination of sizes is that often online advertising campaigns display the same or related advertisements using both these sizes on the same web page. This matched pairing of advertising was used in three of the five sets of advertisements prepared for the websites.

Google Adwords have become extremely common on the web, especially on blogs and smaller websites attempting to monetise website traffic. There are numerous Adword formats, one common format is a 3 x 2 layout at the top or bottom of the main content area. A unit of this sort was constructed for the website, occupying a 550 x 240 pixel space.

### 3.4.3 Advertisement Characteristics

In keeping with the purposes of the study, five sets of advertisements were selected that expressed the following concepts or themes, either aesthetically or through the associated nature or implication of the message or product being promoted:

**Disreputable Subject**
Advertisements with socially dubious or potentially negative associations, e.g. online gambling, obtaining wives online, get rich quick schemes.

**Irrelevant Subject**
Advertisements displaying unrelated content and services such as weight loss or teeth whitening on inappropriate websites e.g technology blogs, news aggregators.

**Distracting Animation**
Numerous examples of animations existed, the most notable being a class of “bait and switch” advertisements with flashing texts like “Congratulations you are the 1,000,000 winner!” or “Warning, spyware detected!”
Ugly Aesthetics

In keeping with the criteria used to determine ugliness in website design, numerous examples of ugly ads were assessed, many of which also combined the irrelevant or disreputable themes outlined above. The final five sets of advertisements chosen were as follows:

**Online Casino Gambling and Date Hot Russian Girls**

![Illustration 20: Online Casino Gambling and Date Hot Russian Girls advertisements](image)

This pair was chosen to best represent the most socially negative options, with the addition of borderline sexual content and general ugliness of the Date Hot Russian Girls ad.

**Teeth whitening matched pair**

![Illustration 21: Teeth whitening advertisements](image)

This pair was chosen for their generally unpleasant closeup anatomical appearance and doubtful claims like “Don't let the dentist fool you”.
“No Diet” Weight loss matched pair

Illustration 22: “No Diet” Weight loss advertisements

Again, an implausible and ugly combination with highly doubtful claims of easy weight loss without dieting or exercise. These ads are at once ugly and disreputable.

Free Online Gaming matched pair

Illustration 23: Free Online Gaming advertisements

Arguably the most innocuous message, this matched pair was chosen mainly for the dominant colouration and animations, although it is still essentially an irrelevant message in the context of an online business directory.

“Congratulations, You are a Winner” and “Google pays me $129 an hour”
These two ads were paired up as classic “bait and switch” examples where the final message on arrival at the advertisers website has little no connection to the promises in the advertisements. The wildly animated “Congratulations” leaderboard was exceptionally irritating to look at, but this quality cannot be effectively communicated in print. Numerous examples of this sort abound on the web however.

The five Adword blocks were all visually similar in layout, but mixed the amounts of relevant and irrelevant subjects, offers and themes to produce unique combinations.

3.4.4 Placement of advertising in the Designs

These advertisements were placed in the study websites in very typical locations. The leaderboard was placed into the navigation bar, which necessitated producing slightly smaller versions of the logos for these website variants. But otherwise, the advertisements were simply added into the design without changing the existing placements.

The Medium Rectangle was introduced into the sidebar of the page, a typical location, and the Adwords block was placed above the main content, again a placement seen in many example directory websites as well as numerous blogs. These placements were identical in both the Attractive and Unattractive designs.

3.5 The Four Completed Designs

The four completed designs are shown in Illustration 25, using the same service and advert combination across the examples.
Research Design and Methodology
Attractive - with Ads

Unattractive - With Ads

Illustration 25: Four final web page designs
3.6 Research Instrumentation

3.6.1 Survey Form

Each service web page contained a hidden survey form that was revealed by clicking a red button at the base of the page in the browser. This form, when revealed, used standard browser form elements on a neutral light grey ground so as not to compete with the styling and colours of the web page content itself. The presentation of this form was identical in all versions of the site.

None of the nine questions could be left blank or unanswered for the form to submit successfully; if any questions were missed, the page redisplayed showing the missed question highlighted. Once the form successfully submitted, the user was returned to an instructions/navigation page.

Importantly, the survey form, when it was visible, could in no way obstruct the ability of the participant to scroll around the page and study aspects of the design while answering the questions. This requirement meant that a common online survey method, the popup window, was not going to be applicable. Popups obscure the content of the page, can easily be accidentally dismissed or lost and sometimes interfere with the ability to scroll. Similarly, opening a new page, or a new tab in the browser, was likely to result in user confusion in a high proportion as cases, especially since there was no simple means of providing oversight or assistance to users having difficulties online.
Ideally, research of this sort would perhaps have been carried out with the survey instrument separate from the material the participants were assessing. A physical paper survey form or interview format might have been ideal, as has been done in other instances in this subject area (Alsudani & Casey, 2009). This was however infeasible, available resources for the research dictated an online survey, so the survey instrument and the survey subject needed to be presented via the same browser window in a manner that did not interfere with the intended function of either aspect.

3.6.2 Quantitative Measures - Semantic Differential Scales

The survey contains 8 quantitative questions represented as 7 point semantic differential scales with bi-polar verbal anchors. Each of these was implemented as 7 radio buttons.

These scales were adapted from those used in prior studies on aspects of credibility by Fogg & H. Tseng (1999), Fogg et al (2001), Lavie & Tractinsky (2004), and standard credibility scales found in the Marketing Scales Handbook (Bruner, Hensel, & James, 2001).

These 8 questions were intended to capture quantitative information to rate the services for the four aspects of credibility identified in the literature review: honesty, expertise, predictability and reputation. Each subject targeted was addressed twice, in paired questions. These questions and their verbal anchors are provided in Table 1.

<table>
<thead>
<tr>
<th>Q.1</th>
<th>Honesty</th>
<th>How trustworthy did the service seem?</th>
<th>Not Trustworthy</th>
<th>Highly Trustworthy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q.2</td>
<td>Expertise</td>
<td>How competent did the service seem?</td>
<td>Incompetent</td>
<td>Highly Competent</td>
</tr>
<tr>
<td>Q.3</td>
<td>Predictability</td>
<td>How reliable did the service seem?</td>
<td>Unreliable</td>
<td>Highly Reliable</td>
</tr>
<tr>
<td>Q.4</td>
<td>Reputation</td>
<td>How successful does the service appear to have been so far?</td>
<td>Unsuccessful</td>
<td>Highly Successful</td>
</tr>
<tr>
<td>Q.5</td>
<td>Honesty</td>
<td>How credible or believable was the information the service gave?</td>
<td>Not Credible</td>
<td>Highly Credible</td>
</tr>
<tr>
<td>Q.6</td>
<td>Expertise</td>
<td>How experienced or knowledgable did the service appear to be?</td>
<td>Inexperienced</td>
<td>Highly Experienced</td>
</tr>
<tr>
<td>Q.7</td>
<td>Predictability</td>
<td>How consistent or thorough was the information the service gave?</td>
<td>Inconsistent</td>
<td>Highly Consistent</td>
</tr>
<tr>
<td>Q.8</td>
<td>Reputation</td>
<td>How highly would you recommend this service to a friend?</td>
<td>Would not recommend</td>
<td>Highly Recommend</td>
</tr>
</tbody>
</table>

Table 1: Survey Questions
These are all highly abstract concepts, and have a significant degree of overlap in their perceived meaning as discussed in the literature review.

Synonyms for the appropriate subjects where possible were built into the questions, and having two questions for each subject was intended to balance errors that might otherwise occur due to individuals personal semantic interpretations.

The second question in each pair was intended to include more concrete terms to compliment the abstract single term in the first question. The exception to this arrangement was questions 4 and 8, which attempt to address both sides of the reputation issue by encouraging the participant to think about how highly others may have considered the service, then to think about how highly they would consider it to others.

### 3.6.3 Qualitative Measures - Free form comment

In addition to these scales, a ninth free-form question was provided as a comments text field, intended to capture emergent qualitative themes. This framing of this question was:

**What is your overall assessment or opinion of this service?**

Placing this question at the bottom of the survey after the semantic differential scales meant that participants had an opportunity to formulate their thoughts and opinions as they resolved their answers to the previous eight questions, and could discuss any key or unresolved issues that had emerged from the process in an unconstrained manner.
4 Validating the Designs – The Pilot Study

One of the primary early goals of this research was to establish a solid set of criteria about what constituted an attractive and unattractive web page design. Some previous studies had approached this problem obliquely, for example Harper et al (2009) declared their choices to represent a “good spread of standard visual types” based on “an understanding of the principles of design” (p.9). Others relied on heuristics (Sutcliffe, 2001), expert evaluations and analysis (Park et al., 2004) to identify example material for study by non-expert participants that meet certain design criteria.

Using a pilot study (Kurosu & Kashimura, 1995) or direct empirical testing of user opinions on attractiveness of a design (Tractinsky et al., 2000) is an established means of determining a broader consensus, at least within a constrained set of options. So a pilot study was implemented to seek confirmation that the specific design features chosen to distinguish the Attractive and Unattractive directories for this study were generally having the right effect.

4.1 Pilot Study Design

4.1.1 Study of Design Features and Ranking

A pilot study was created to ascertain that the two basic designs do appear distinct and that one indeed seems more unattractive than the other. Additionally, the pilot was used to establish overall placement of these designs relative to a subset of the example set of 12 reference websites used in the creation of these designs.

The study, being a small pilot, was to be conducted on a face to face basis, using a Mac laptop to present the on screen component, and paper questionnaires to gather the age, gender, internet experience and answers of the participants. An example questionnaire is included in Appendix 2.

Versions of the same fictional service, “Drain Master Plumbing” were made for both the Attractive and Unattractive versions of the website. To make these web pages more plausible as actual directory pages in a wider context, they were presented with panels of adwords text advertising incorporated in the design, but with no banner ads. This inclusion of some text ads meant the designs had much more similar overall content to the sample designs that had originated in real online directories, rather than representing idealised extremes. Illustration 27 shows these designs as they appeared in the Pilot Study.
Seven of the reference websites originally identified were also linked to directly from the navigation page, as seen in Illustration 28, as animated advertising was present in a number of the examples and a screenshot would not have represented this feature.

Each participant was interviewed to complete the personal data section of the questionnaire to gather their age, gender and internet experience. Then the participants were shown each of the 9 screens in a random order.

Once they had viewed the screen for a minute or so, they were asked to complete a questionnaire about that screen, which contained a group of 3 semantic differential scale questions (9 point scale) with bi-polar verbal anchors as seen in Table 2. There were two additional freeform questions about their thoughts regarding attractive features on the page.

### Table 2: Pilot Study Questions

<table>
<thead>
<tr>
<th>Q1</th>
<th>Overall Reactions to the page 1</th>
<th>Terrible - Wonderful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2</td>
<td>Overall Reactions to the page 2</td>
<td>Ugly - Beautiful</td>
</tr>
<tr>
<td>Q3</td>
<td>Overall Reactions to the page 3</td>
<td>Dull - Stimulating</td>
</tr>
<tr>
<td>Q4</td>
<td>What was the most attractive feature?</td>
<td>Freeform</td>
</tr>
<tr>
<td>Q5</td>
<td>What was the most unattractive feature?</td>
<td>Freeform</td>
</tr>
</tbody>
</table>
Once they had completed the questionnaires for all 9 screens, the participants were asked to order printed copies of the screens from “worst” to “best” based purely on how they looked, the aesthetics of the designs. This sorting methodology is a common data-gathering technique for a variety of cognitive and perceptual issues, and had been used in previous studies around aesthetics and credibility (J. Kim & Moon, 1998).

### 4.1.2 Study of Advertising Ranking

A second component of the pilot study involved displaying the same Attractive version of the “Drain Master Plumbing” service with the 5 different combinations of banner ads that were going to be used in the main study. The participants were asked to examine these 5 screens,
seen in Illustration 29, then take printed copies of the screens and again stack them according to their own assessment of the “worst” to “best”, based on their feelings or reactions to the advertisements, as well as how they looked.

Illustration 29: Advertising Pilot Screens

4.2 Pilot Study Participation

9 participants took part in the pilot, 5 men and 4 women aged between 18 and 45, all associates, workmates and neighbours of the researcher. The background knowledge of each participant was very different. Most used the Internet on a daily basis and some used it frequently. Participants in the pilot study were specifically questioned on the appearance of the websites themselves, and were therefore not involved in the main studies.

4.3 Comparative Website Design Results

4.3.1 Semantic Differential Question Results

The three sets of questions using semantic differential scales were scored from 0 to 9, and each attempted to examine a different aspect of the participants affective attitude toward the designs they were examining. The results appear in the tables below.
The first question, scoring the designs on a scale between Terrible and Wonderful, saw the two looksie designs at opposite ends of the scale, with the Unattractive design at the bottom.

Table 3: Pilot Study results for Terrible - Wonderful question (ranked by mean)

<table>
<thead>
<tr>
<th>Webpage Design</th>
<th>Avg. Score</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shop NZ</td>
<td>Mean</td>
<td>5.33</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>3.250</td>
</tr>
<tr>
<td>Looksie - Attractive</td>
<td>Mean</td>
<td>5.11</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>1.361</td>
</tr>
<tr>
<td>Finda</td>
<td>Mean</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>3.500</td>
</tr>
<tr>
<td>Scratch &amp; Save</td>
<td>Mean</td>
<td>4.56</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>3.778</td>
</tr>
<tr>
<td>Zipleaf</td>
<td>Mean</td>
<td>4.56</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>5.278</td>
</tr>
<tr>
<td>Topmaq</td>
<td>Mean</td>
<td>4.11</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>2.611</td>
</tr>
<tr>
<td>TamilBizCard</td>
<td>Mean</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>5.278</td>
</tr>
<tr>
<td>Prime Imports</td>
<td>Mean</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>5.194</td>
</tr>
<tr>
<td>Looksie - Unattractive</td>
<td>Mean</td>
<td>3.22</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>2.694</td>
</tr>
</tbody>
</table>

The second question, scoring the designs on a scale between Dull and Stimulating, again saw the Unattractive design at the bottom, although the Attractive design slipped to third.

Table 4: Pilot Study results for Dull - Stimulating question (ranked by mean)

<table>
<thead>
<tr>
<th>Webpage Design</th>
<th>Avg. Score</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shop NZ</td>
<td>Mean</td>
<td>5.67</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>1.750</td>
</tr>
<tr>
<td>Topmaq</td>
<td>Mean</td>
<td>4.78</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>5.194</td>
</tr>
<tr>
<td>Looksie - Attractive</td>
<td>Mean</td>
<td>4.67</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>2.250</td>
</tr>
<tr>
<td>Zipleaf</td>
<td>Mean</td>
<td>4.44</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>1.778</td>
</tr>
<tr>
<td>Finda</td>
<td>Mean</td>
<td>4.33</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>3.250</td>
</tr>
</tbody>
</table>
The third question, scoring the designs on a scale between Ugly and Beautiful, saw the Unattractive design again at the bottom, and the Attractive design in second place.

### Table 5: Pilot Study results for Ugly - Beautiful question (ranked by mean)

<table>
<thead>
<tr>
<th>Webpage Design</th>
<th>Avg. Score</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shop NZ</td>
<td>Mean</td>
<td>5.89</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>3.11</td>
</tr>
<tr>
<td><strong>Looksie - Attractive</strong></td>
<td><strong>Mean</strong></td>
<td><strong>5.11</strong></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>1.86</td>
</tr>
<tr>
<td>Zipleaf</td>
<td>Mean</td>
<td>5.11</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>1.86</td>
</tr>
<tr>
<td>Scratch &amp; Save</td>
<td>Mean</td>
<td>4.89</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>4.11</td>
</tr>
<tr>
<td>Finda</td>
<td>Mean</td>
<td>4.22</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>2.19</td>
</tr>
<tr>
<td>TamilBizCard</td>
<td>Mean</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>4.78</td>
</tr>
<tr>
<td>Prime Imports</td>
<td>Mean</td>
<td>3.44</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>4.03</td>
</tr>
<tr>
<td>Topmaq</td>
<td>Mean</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>4.75</td>
</tr>
<tr>
<td><strong>Looksie - Unattractive</strong></td>
<td><strong>Mean</strong></td>
<td><strong>3.00</strong></td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>2.75</td>
</tr>
</tbody>
</table>

#### 4.3.2 Freeform comments

Comments made in the two opposing freeform questions were analysed and coded for words and phrases associated with positive or negative themes, terms and concepts. The results of this
Table 6: Pilot Study results website * Direction of Comment cross tabulation

<table>
<thead>
<tr>
<th>Website</th>
<th>Direction</th>
<th>Count</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>Positive</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Looksie - Attractive</td>
<td>Count</td>
<td>10</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>19.2%</td>
<td>80.8%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>5.1%</td>
<td>16.8%</td>
</tr>
<tr>
<td>Finda</td>
<td>Count</td>
<td>18</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>31.6%</td>
<td>68.4%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>9.2%</td>
<td>15.6%</td>
</tr>
<tr>
<td>Scratch and Save</td>
<td>Count</td>
<td>21</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>39.6%</td>
<td>60.4%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>10.8%</td>
<td>12.8%</td>
</tr>
<tr>
<td>Shop NZ</td>
<td>Count</td>
<td>19</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>40.4%</td>
<td>59.6%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>9.7%</td>
<td>11.2%</td>
</tr>
<tr>
<td>Zipleaf</td>
<td>Count</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>45.8%</td>
<td>54.2%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>11.3%</td>
<td>10.4%</td>
</tr>
<tr>
<td>Tamil Biz Card</td>
<td>Count</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>47.7%</td>
<td>52.3%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>10.8%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Topmaq</td>
<td>Count</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>46.5%</td>
<td>53.5%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>10.3%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Looksie - Unattractive</td>
<td>Count</td>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>54.8%</td>
<td>45.2%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>11.8%</td>
<td>7.6%</td>
</tr>
<tr>
<td>Prime Imports</td>
<td>Count</td>
<td>41</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>69.5%</td>
<td>30.5%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>21.0%</td>
<td>7.2%</td>
</tr>
<tr>
<td>Total</td>
<td>Count</td>
<td>195</td>
<td>250</td>
</tr>
<tr>
<td></td>
<td>% within website</td>
<td>43.8%</td>
<td>56.2%</td>
</tr>
<tr>
<td></td>
<td>% within Direction</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

4.3.3 Manual Ranking Results

The participants ranked printed versions of the website designs after completing the questions, giving them an opportunity to reconsider their overall judgement having seen all options. This
result largely supported the questionnaire results, with the Attractive version coming fourth and the Unattractive coming last again as seen in Table 7.

### Table 7: Pilot study results - web page design ranking

<table>
<thead>
<tr>
<th>Web page design</th>
<th>Avg. Rank</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shop NZ</td>
<td>3.00</td>
<td>.957</td>
</tr>
<tr>
<td>Finda</td>
<td>3.11</td>
<td>.716</td>
</tr>
<tr>
<td>Zipleaf</td>
<td>3.89</td>
<td>.790</td>
</tr>
<tr>
<td><strong>L ooksie Attractive</strong></td>
<td><strong>4.22</strong></td>
<td><strong>.572</strong></td>
</tr>
<tr>
<td>Scratch and Save</td>
<td>4.78</td>
<td>.778</td>
</tr>
<tr>
<td>Prime Imports</td>
<td>6.00</td>
<td>.986</td>
</tr>
<tr>
<td>Topmaq</td>
<td>6.22</td>
<td>.830</td>
</tr>
<tr>
<td>TamilBizCard</td>
<td>6.67</td>
<td>.500</td>
</tr>
<tr>
<td><strong>L ooksie Unattractive</strong></td>
<td><strong>7.11</strong></td>
<td><strong>.455</strong></td>
</tr>
</tbody>
</table>

### 4.3.4 Conclusions of the Website Design Component

Overall, there was a clear distinction between the perceived attractiveness of the two designs, with the Unattractive design coming last in almost all the measures made, and the Attractive design coming in the upper third. This result, although obviously with too small a sample size to be significant in any way, confirmed the suitability of the two designs to be used for the main studies. Interestingly, it was a positive sign that the Attractive version did not always appear as most attractive in all measures, as the intention was always to achieve a plausible baseline rather than a pinnacle of design virtue.
4.4 Comparative Advertising Results

The advertising ranking was a much simpler measurement, basically consisting of participants viewing the 5 designs onscreen to assess the animation effects, then ranking printouts of the 5 variations manually. Participants placed the “Congratulations, You are a Winner” and “Google pays me $129 an hour” option at the bottom, and considered the “Brainbend Free Online Gaming” matched pair the most innocuous as seen in Table 8. This result was not surprising as the Brainbend pair, even though animated and certainly irrelevant to the context of the directory pages in which the ads were seen, arguably had the least objectionable or disreputable content.

Table 8: Pilot Study results - Advertisement Ranking

<table>
<thead>
<tr>
<th>Advert</th>
<th>Mean</th>
<th>Variance</th>
<th>Avg. Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brainbend</td>
<td>Mean</td>
<td>.861</td>
<td>1.89</td>
</tr>
<tr>
<td>Bleach</td>
<td>Mean</td>
<td>1.694</td>
<td>2.78</td>
</tr>
<tr>
<td>Casino</td>
<td>Mean</td>
<td>2.750</td>
<td>3.00</td>
</tr>
<tr>
<td>Diet</td>
<td>Mean</td>
<td>1.500</td>
<td>3.33</td>
</tr>
<tr>
<td>Congratulations</td>
<td>Mean</td>
<td>1.750</td>
<td>4.00</td>
</tr>
</tbody>
</table>

The violently animated and highly dubious offers of the “Congratulations, You are a Winner “ pair were clearly the worst received in this small sample. There was also an interesting but predictable gender split around the Online Casino Gambling and Date Hot Russian Girls option, with men preferring it more than women, producing a far wider variance than the other ads, although overall it appeared near the middle.

4.4.1 Conclusions of the Advertising Component

Overall, the more socially objectionable, untrustworthy or disreputable the advertisement message was, the worse it performed. The dubious, too good to be true claims of high speed diets and unlikely financial rewards appeared at the bottom. Animated ads appeared at both the top and bottom positions in this small sample, suggesting that the role of animation in the participants affective assessments was unclear. The role of the general aesthetics of the ads was also unclear.
4.5 Conclusions and modifications to the Website Designs

For the purposes of determining if the two web page designs were sufficiently different and whether they were attractive or not, this study confirmed that the basic designs were promising.

In some individual cases there was less obvious distinction made between the Attractive and the Unattractive design than was expected, although only one participant actually preferred the Unattractive to the Attractive. Although the majority of participants ranked the Attractive version in the top third, and the Unattractive version in the lower third, the overall rankings of individual designs from amongst the reference designs was not as consistent as had been expected.

A common comment theme raised against the Attractive design was that it was bland or average and several other designs were positively identified as being bright. In response to this, the Attractive design colour scheme was made stronger, the blue and green more saturated and with increased contrast, and the search bar was given a more dimensional and stronger colour scheme.

The Unattractive design was left essentially unchanged, except that the font size was reduced as small fonts had been identified as negative several times, and some more misalignment was added to the search bar region to add to the overall sense of disorder.
5 Between Groups Research

5.1 Between Groups Research Overview

A 2 x 2 between groups design with two independent variables was implemented to explore the effects of Aesthetic differences and the presence of irrelevant Advertising on user perceptions of the credibility of services seen on an Online Directory.

The dependent variable for this research was:

\[ \text{DV. The Credibility of services depicted on the website.} \]

The independent variables for this research were:

\[ \begin{align*}
\text{IV1. Website Design Attractiveness} \\
\text{IV2. The presence of Advertising of low relevance}
\end{align*} \]

Each of these variables explores one of the hypotheses \( H_1 \) and \( H_2 \) described above.

A key reason for using a between groups design, despite its lower statistical power, was to explore whether the effect of these variables was significant even when they were not seen in comparison with any other different context, ie. were the effects of the two variables somewhat absolute and persistent or were they only perceived relative to other different designs experienced in close proximity of time or space.

To help this process, the experiment was intended to be as similar to the experience of using an actual online business directory as feasible, so that any effects of the variables would be reinforced by multiple exposure on numerous similar pages.

5.2 Between Groups Instrument Design

5.2.1 Simulating an Online Directory

The full set of 16 fictional Auckland business services were employed in this between groups version of the study. There were four service types:

- Accountant
- Plumber
with four examples of each, all derived from real, anonymised web content as explained above. The intention of having so many services on the website was to help produce a more complete and convincing impression of a real online directory, and also to give multiple opportunities for the participants to experience the particular features of the design or advertising depicted to produce a more realistic response.

It had also been intended for this design to support the ability to navigate freely between services using a services listing page, so that the users might interact with an even more realistic model of an online directory, and this page represented an additional opportunity for the advertising and aesthetic variables to be presented to the user.

This page had been conceived and implemented as a fictitious search results page, which would have echoed the same Attractive or Unattractive stylings and advertising as the service listing pages. The concept was to increase the credibility of the overall experience as being a valid website, to give more opportunities for affective impression of the website to be built up.

This page was the landing page for commencement of the survey and also the page that the user automatically returned to after submitting each survey form.

Unfortunately, a number of observed users tests quickly showed that participants arrived at this page having little or no idea what they were to do next. The surveys were only accessible from the listing pages, and any instructions given as part of the consent process were no longer visible to aid participants in seeking out these survey’s or even knowing they should click a link to view a service, despite the fact that the search results/detail page information architecture has been the dominant method of using directory and e-commerce websites for many years.

Since this was not a usability test, and there was no feasible way to include instructions in the content of this page without undermining the aesthetic impression in ways too complex to predict, the page was redesigned to be an information and navigation tool rather than visually part of the experiment.

The page was provided with instructions and effectively became the home page where participants arrived and returned to on the viewing and submission of each survey form. This page was neutrally styled, in the same manner as the introductory pages, as seen in Illustration 30.
Even once this was implemented, user tests showed the need for progress indicators and greying out of completed services to help users progress freely to the relevant services.

Importantly, as discussed in the limitations later, this method of presenting the services was not properly counterbalanced, which meant that comparisons between services and advertising types, while not specifically required by the hypotheses, was not going to be possible. This design flaw was corrected in the subsequent study.

5.2.2 Functionality of the Website

It should be noted that this website was not fully functional in any sense. Where normally in an online directory there would be a search engine and a search results page that linked to individual details pages across a wide range of topics, only the 16 details pages specific to the four types of services were created for this research. In a sense only the ‘front end’ of the site was usable: any ‘deeper’ linking to other pages or parts of the site was not supported, the links to these pages still appeared as a valid part of the overall design, but the links did nothing.

This is a common approach seen within the web design industry when presenting working prototypes to a client. The purpose is to provide a valid overview of the site’s general visual layout and the navigational structure of the interface. In this case it was essential to support all the apparent aesthetic components of a web page visually, without going to the extent of implementing the whole website.
### 5.3 Between Groups Research Results

#### 5.3.1 Participants & Recruitment

A total of 282 participants were recruited for this Web-based study via personal and professional contacts on LinkedIn and Facebook, email announcements to Massey University MBA classes and posters displayed at the Massey University Albany Campus in January 2011. The participants were all self-selected, unscreened and participated voluntarily. A movie voucher was offered on completion of the research by way of compensation or “koha” for their time and effort spent on the research.

Of the 282 participants recruited, 132 complete usable sets of data were collected, with a minimum of 28 subjects per website version. Only 132 participants completed the entire survey (possibly because the survey was quite long). Each participant was exposed to only one website version, and assignment of participants to versions was randomised.

The only personal details captured in the survey about the participants was their age, sex and internet experience. The potential participants were assured that no identifying information would be required to take or even finish the survey, but that address details could be volunteered at the end to enable compensation to be mailed out to them.

The final sample displayed a fairly even split between males and females (57 males to 74 females). Ages ranged from 18 to 65 + with 84.1% of participants aged between 25 and 44, and 50% aged between 35 and 44 years old. 97% of participants used the internet most days and 64.3% many times a day.

All participants accessed the same website to participate, and before beginning the survey participants read an information page to be aware of risks, benefits, alternatives, confidentiality, payments, contact information, and the voluntary nature of participating in the study.

A single paragraph on this page outlined the general premise of the research, its investigation into attitudes and impressions formed, but deliberately made no specific mention of aesthetics or advertising:

> “This research explores the perceptions and opinions that participants form about companies and individuals that are represented on a fictitious business directory website. A series of short online questionnaires is completed by the participants, and the entire process can be completed online in the participants own time in under one hour”

- from the Information Sheet
Informed consent was required before participating in the Web survey. Only participants that acknowledged their agreement with the consent form via a checkbox were allowed to take the Web survey. Screenshots of these pages can be found in Appendix 3.

Once informed consent was given, participants found themselves on the instruction and navigation page mentioned earlier that gave simple instructions for the experiment and listed links to the 16 sample website pages. Participants were asked to visit each of the linked pages, study it for a minute or two, then click the red button on the bottom of that page to complete the questionnaire. After visiting a given webpage and completing its survey, the participant found themselves back at this instruction and navigation page automatically.

The page had a counter showing the number of completed surveys and the number remaining. Additionally, as each page was viewed and its survey was completed, the link to that page was greyed out and disabled to ensure participants could easily determine which pages they had not yet viewed.

Users controlled their own progression through the web pages, and their online sessions lasted an average of 53.2 minutes. Once they had successfully completed all 16 questionnaires, the participants were directed to a completion page and thanked for their participation. An opportunity was provided at this point for participants to claim a Movie voucher as compensation for their time and trouble.

### 5.4 Analysis

#### 5.4.1 Validating the data

The results of the 8 semantic differential scales were averaged across the 16 services experienced for each participant, since there was no counterbalancing in the experiment, so each participant received a mean value for each of the 8 credibility related questions.

These mean results were plotted and frequencies tabulated to test for normal distribution, and to detect any skew and kurtosis (or ‘peakedness’) of the data as seen in Illustration 31.

Z-scores, a form of standardised distribution where all values are between 0 and 1 were calculated from these results and calculations made for skewness and kurtosis which showed that 7 out of the 8 sets of results had no significant skew or kurtosis, as none had an absolute value greater than 1.96, which is the threshold of significance at $p < .05$. These results are in Table 9.
Q.8, the reputation question “How highly would you recommend this service to a friend?” was significantly negatively skewed at -2.398, implying lower scores overall for this question. Kurtosis for Q.8 was not significant.

### Table 9: Skewness and Kurtosis results

<table>
<thead>
<tr>
<th></th>
<th>Q.1 honesty</th>
<th>Q.2 expertise</th>
<th>Q.3 predict</th>
<th>Q.4 reputation</th>
<th>Q.5 honesty</th>
<th>Q.6 expertise</th>
<th>Q.7 predict</th>
<th>Q.8 reputation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>132</td>
<td>132</td>
<td>132</td>
<td>132</td>
<td>132</td>
<td>132</td>
<td>132</td>
<td>132</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.586</td>
<td>.586</td>
<td>.586</td>
<td>.545</td>
<td>.475</td>
<td>.570</td>
<td>.543</td>
<td>.934</td>
</tr>
<tr>
<td>Skewness</td>
<td>-.021</td>
<td>.065</td>
<td>.110</td>
<td>.113</td>
<td>.138</td>
<td>.099</td>
<td>.118</td>
<td>-.506</td>
</tr>
<tr>
<td>Std. Error of Skewness</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
<td>.211</td>
</tr>
<tr>
<td>Skewness z-score</td>
<td>-.099</td>
<td>.308</td>
<td>.521</td>
<td>.535</td>
<td>.654</td>
<td>.469</td>
<td>.559</td>
<td>-.2398</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-.254</td>
<td>-.185</td>
<td>-.194</td>
<td>-.270</td>
<td>-.382</td>
<td>-.308</td>
<td>-.388</td>
<td>.152</td>
</tr>
<tr>
<td>Std. Error of Kurtosis</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
<td>.419</td>
</tr>
<tr>
<td>Kurtosis z-score</td>
<td>-.606</td>
<td>-.442</td>
<td>-.463</td>
<td>-.644</td>
<td>-.912</td>
<td>-.735</td>
<td>-.926</td>
<td>.363</td>
</tr>
</tbody>
</table>
Additionally, Kolmogorov-Smirnov (K-S) tests were run to establish that the distribution was normal and all 8 questions returned no significant normality issues as follows:

Q.1 Honesty, D(132) = 0.053, ns
Q.2 Expertise, D(132) = 0.040 ns
Q.3 Predictability, D(132) = 0.045, ns
Q.4 Reputation, D(132) = 0.054, ns
Q.5 Honesty, D(132) = 0.064, ns
Q.6 Expertise, D(132) = 0.046 ns
Q.7 Predictability, D(132) = 0.047, ns
Q.8 Reputation, D(132) = 0.062, ns

Again the most non-normal, but still not significant result was for Q.8 “How highly would you recommend this service to a friend?”

A transformation for the data was considered to address the skewness in Q.8 which threatened the robustness of any intended ANOVA result, but after consideration, the data for Q.8 was dropped.

5.4.2 Validating the survey questions and extracting factors

The dependant variable for the research was credibility, as expressed by aspects of:

- Honesty
- Expertise
- Predictability
- Reputation

There were two questions for each term, derived from previous studies as outlined before, with the exception of reputation which now only had one question. It was important to establish that the remaining seven questions did not introduce other unrelated factors.

Factorial Analysis by Principle Component Analysis (PCA) was used to determine if the seven questions were in fact all exploring related parts of the same dimension, and that there was high correlation between them.

The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .952 (‘superb’ according to Field (2009)). Bartlett’s test of sphericity $X^2 (21) = 2144.55$, $p < .001$, indicated that the correlations between items were sufficiently large for PCA.
Only one component or factor with a eigenvalue greater than Kaiser's criterion of 1 was identified, which explained 95.25% of the variance, a result which is consistent with expectations of the four aspects of credibility explored each being an integral part of a single factor.

A reliability analysis was also done, showing that a single factor consisting of Q.1 - Q.7 had a high reliability with Cronbach's $\rho = .992$. Since all the questions were so highly correlated, the mean of all 7 questions was used as the basis for the new factor, Credibility which was used for further analysis.

5.4.3 Analysis with Two-Way Independent Factorial Anova

A two-way independent factorial Anova was performed using the dependent variable Credibility derived from the factor analysis, and the two independent variables Appearance and Presence of Advertising to determine if there was support for:

H1. Reduced website attractiveness negatively affects users perceptions of the credibility of the services represented by the directory.

H2. The presence of website advertising of low relevance negatively affects users perceptions of the services represented by the directory.

Homogeneity of Variance, a measure of the variance between groups in the research, was tested with Levene's test which was found not significant at $F(3,128) = .867$, ns.

This ANOVA revealed that there was a significant main effect in the predicted direction of the Presence of Advertising on the perceived Credibility of the services presented on the online directory website $F(1,128) = 16.6, p < .001, \quad ^2 = .1$, supporting H2, as seen in Table 10.

<table>
<thead>
<tr>
<th>Table 10: Two way Factorial ANOVA - tests of Between Groups effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
</tr>
<tr>
<td>appearance</td>
</tr>
<tr>
<td>advertising</td>
</tr>
<tr>
<td>appearance * advertising</td>
</tr>
<tr>
<td>error</td>
</tr>
</tbody>
</table>

There was however a non-significant main effect of Appearance on the perceived Credibility of the services presented on the online directory website $F(1,128) = 1.68, p = .197, \quad ^2 = .004$. This

Between Groups Research 85
was an unexpected and surprising result, particularly since the trend, though not significant, was in the opposite direction from that anticipated as seen in Illustration 32:

Additionally, the interaction effect between Appearance and Presence of Advertising was not significant at $F(1,128) = .04, p = .841$, $\hat{\eta}^2 = .004$, but it again trended in the opposite direction to that anticipated. Specifically, the credibility was more similar for Attractive Appearance without Advertising ($M = 4.77$, $SD = .692$) vs Unattractive Appearance without Advertising ($M = 4.90$, $SD = .549$) than for Attractive Appearance with Advertising ($M = 4.25$, $SD = .736$) vs Unattractive Appearance with Advertising ($M = 4.43$, $SD = .718$)

While overall not a significant result, it arguably implies that at least in this between groups context, the Presence of advertising was found slightly more acceptable on the Unattractive site than on the Attractive.

Consequently, at least within the context of this between groups experiment, we must accept the null hypothesis for $H1$. 

Illustration 32: Profile Plot of Appearance * Advertising

![Illustration 32: Profile Plot of Appearance * Advertising](image-url)
5.5 Limitations of this study

Between groups designs have intrinsically less statistical power to determine significance than within subjects designs, but the choice of this methodology was made as part of an attempt to more closely reproduce the fundamental characteristics of using different Online Directories, which are usually consistent in their strengths and weaknesses and do not alternately show different aesthetic designs or different degrees of advertising. Including 16 different services in the design was partly to reproduce this more immersive directory experience, but it was also intended to allow examination of effects on different service types and to explore the effect of different advertising types.

As a consequence of this design decision, and some late stage changes in response to usage studies during the pilot testing, there was an error in the design that diminished the capability of the study to examine some finer grained details. The presentation of services was left to the participants discretion, and was not counterbalanced. This lack of counterbalancing meant that examining the effects on individual services or different types of advertising was not feasible. The design was however still capable of addressing the first two hypotheses, and H2 was supported by the results and analysis.

In response to these limitations a second within subjects design was devised that had fewer services but was fully counterbalanced in its presentation to participants, and otherwise used the same research instrument.
6 Within Subjects Research

6.1 Within Subjects Research Objectives

A 2 x 3 within subjects design with two independent variables was implemented to explore the effects of Aesthetic differences and the presence of Advertising on user perceptions of the credibility of services seen on an Online Directory.

The dependent variable for this research was:

DV. The Credibility of services depicted on the website.

The independent variables for this research were:

IV1. Website Design Attractiveness
IV2. The presence of Advertising of low relevance

IV2 had three levels in this experiment, No Adverts, Irrelevant Adverts and Disreputable Adverts. These second two types were selected from the range of advertisements used in the first study, with the Brainbend Games advertisement taking on the role of irrelevant advert, and the Hot Russian Brides advertisement taking the role of disreputable advert.

These were chosen after consideration of the free form feedback from the first study, which is analysed in section 7. The Hot Russian advertisement had engendered the most negative feedback with many comments having overtones of disrepute, disgust and general lack of appeal.

This additional level of advertising was added to enable the exploration of H3 as well as supporting the existing capability of addressing hypotheses H1 and H2 that was also in the original design.

6.2 Within Subjects Research Instrument

6.2.1 Counterbalancing

A reduced set of 6 fictional Auckland business services from the original study were employed in this within subjects version. There were three service types:
Accountant
Restaurant
Hairdresser

with two examples of each, each derived from the same real, anonymised web content as before.

These 6 services were fully counterbalanced for order of presentation and combination of Appearance and Advertising, and unlike the previous study, the progression of the participant through their session was automated rather than using a listing page.

In every other way, including the number of questions used, their presentation on the web pages and the key information on the introductory and consent pages, this version of the instrument was identical to the first.

6.3 Within Subjects Research Results

6.3.1 Participants & Recruitment

A total of 151 participants were recruited for this Web-based study again via personal and professional contacts on Facebook, and posters displayed at the Massey University Albany Campus in February 2011. The participants were all self selected, unscreened and participated voluntarily. A movie voucher was again offered on completion of the research by way of compensation or “koha” for their time and effort spent on the research, even though the survey duration was far shorter for this version, less than 20 minutes.

Of the 151 participants recruited, 121 complete usable sets of data were collected. Each participant was exposed to all 6 versions of the website, and assignment of participants was fully randomised and counterbalanced to control for unwanted practice or boredom effects.

Again the only personal details captured in the survey about the participants was their age, sex and internet experience, although again address details were usually volunteered at the end to enable compensation to be mailed out to them.

The final sample displayed a less even split between males and females (49 males to 71 females) than the first experiment. Ages ranged from 18 to 65 + with 62.8% of participants aged between 25 and 44, and 35% aged between 18 and 25 years old, mainly due to the higher proportion of students who participated as the survey ran during the beginning of first semester rather than during the summer school as had been the case in the first survey. However this set of
participants were heavy internet users, as 97.5% of participants used the internet most days, 82.7% using it many times a day.

As in the previous study, all participants accessed the same website to participate, and before beginning the survey participants read an information page to be aware of risks, benefits, alternatives, confidentiality, payments, contact information, and the voluntary nature of participating in the study.

As before, the participants were asked to study each page for a minute or two, then click the red button on the bottom of that page to complete the questionnaire. However, this study used an automated mechanism that directed the users to a new page when the survey form was submitted until all 6 services had been viewed. Being a much shorter survey than the previous between groups version, the average participant duration was 25.4 minutes.

6.4 Analysis

6.4.1 Validating the Data

A large sample was gathered from results of the 8 semantic differential scales for each of the 6 services the participants studied. This data was plotted and frequencies tabulated to test for normal distribution, and to detect any skew and kurtosis of the data as seen in Illustration 33.

![Histograms of 8 semantic differential results](image)

**Illustration 33: Histograms of 8 semantic differential results**

Within Subjects Research
Z-scores were calculated from these results and calculations made for skewness and kurtosis which showed that, although both skew and particularly kurtosis were much higher than the previous study, as seen in Table 11, the large sample size of 726 renders the result less significant due to the Central Limit theorem. The skew on Q.8 was again higher than the other questions however.

Table 11: Skewness and Kurtosis results

<table>
<thead>
<tr>
<th></th>
<th>Q.1 honesty</th>
<th>Q.2 expertise</th>
<th>Q.3 predict.</th>
<th>Q.4 reputation</th>
<th>Q.5 honesty</th>
<th>Q.6 expertise</th>
<th>Q.7 predict.</th>
<th>Q.8 reputation</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>726</td>
<td>726</td>
<td>726</td>
<td>726</td>
<td>726</td>
<td>726</td>
<td>726</td>
<td>726</td>
</tr>
<tr>
<td>Mean</td>
<td>4.09</td>
<td>4.19</td>
<td>4.18</td>
<td>4.06</td>
<td>4.02</td>
<td>4.12</td>
<td>4.12</td>
<td>3.55</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.599</td>
<td>1.645</td>
<td>1.635</td>
<td>1.737</td>
<td>1.672</td>
<td>1.677</td>
<td>1.656</td>
<td>1.816</td>
</tr>
<tr>
<td>Variance</td>
<td>2.557</td>
<td>2.706</td>
<td>2.672</td>
<td>3.018</td>
<td>2.795</td>
<td>2.814</td>
<td>2.741</td>
<td>3.298</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.090</td>
<td>-0.091</td>
<td>-0.099</td>
<td>0.018</td>
<td>-0.027</td>
<td>-0.017</td>
<td>-0.017</td>
<td>-0.049</td>
</tr>
<tr>
<td>Std. Error of</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
<td>0.091</td>
</tr>
<tr>
<td>Skewness z-score</td>
<td>-0.989</td>
<td>-1.000</td>
<td>-1.087</td>
<td>0.197</td>
<td>-0.297</td>
<td>-0.187</td>
<td>-0.538</td>
<td>2.044</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.715</td>
<td>-0.895</td>
<td>-0.832</td>
<td>-0.999</td>
<td>-0.859</td>
<td>-0.903</td>
<td>-0.886</td>
<td>-1.042</td>
</tr>
<tr>
<td>Std. Error of</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
<td>0.181</td>
</tr>
</tbody>
</table>

Kolmogorov-Smirnov (K – S) tests were not run on this data as they are unreliable on such large sample sizes and give false positives.

6.4.2 Validating the survey questions and extracting factors

Factorial Analysis by Principle Component analysis (PCA) was used to determine if the eight questions were all exploring related parts of the same dimension as was the case in the first study.

The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .960 ('superb' according to Field (2009)). Bartlett's test of sphericity $X^2 (28) = 7963.48$ $p < .001$, indicated that the correlations between items were sufficiently large for PCA.
Again only one component or factor with a eigenvalue greater than Kaiser's criterion of 1 was identified, which explained 85.22% of the variance, which although less than the previous study, is consistent with expectations of all questions being an integral part of a single factor.

A reliability analysis was also done, showing that a single factor consisting of Q.1 - Q.8 had a high reliability with Cronbach's $= .975$. Since all the questions were so highly correlated, the mean of all 8 questions was again used as the basis for the new factor, Credibility.

6.4.3 Analysis with Two - Way Independent Repeated Measures Anova

A two-way independent Repeated Measures Anova was performed using the dependent variable Credibility derived from the factor analysis, and the two independent variables Appearance and Presence of Advertising to determine if there was support for:

H1. Reduced website attractiveness negatively affects users perceptions of the credibility of the services represented by the directory.

H2. The presence of website advertising of low relevance negatively affects users perceptions of the services represented by the directory.

H3. The presence of website advertising of low relevance and low reputability negatively affects users perceptions of the services represented by the directory more than low relevance alone.

Mauchly's Test indicated that Sphericity, which measures the variances of the differences between levels was not significant with Advertisements $X^2(2) = 2.83$, ns and Advertisements * Appearances $X^2(2) = 0.71$, ns. No result was returned for the main effect of Appearances as it only had two levels so sphericity was not an issue.

This ANOVA, seen in Table 12, revealed that there was a significant main effect of Appearance on the perceived Credibility of the services presented on the website $F(1,120) = 26.69$, $p < .001$, $r = .18$, supporting H1.

The ANOVA also revealed that there was a significant main effect of the Presence of Advertising on the perceived Credibility of the services presented on the website $F(2,240) = 23.47$, $p < .001$, $r = .22$, supporting H2.

The interaction effect between Appearance and Presence of Advertising was not significant at $F(2,240) = 1.43$, $p = .246$. 

Within Subjects Research
Table 12: Two way Repeated Measures ANOVA - tests of Within Subjects effects

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>appearance</td>
<td>.787</td>
<td>1</td>
<td>.787</td>
<td>1.680</td>
<td>.197</td>
</tr>
<tr>
<td>advertising</td>
<td>7.783</td>
<td>1</td>
<td>7.783</td>
<td>16.600</td>
<td>.000</td>
</tr>
<tr>
<td>appearance * advertising</td>
<td>.019</td>
<td>1</td>
<td>.019</td>
<td>.040</td>
<td>.841</td>
</tr>
<tr>
<td>error</td>
<td>60.014</td>
<td>128</td>
<td>.469</td>
<td>.469</td>
<td>.469</td>
</tr>
</tbody>
</table>

Additionally, Helmert’s contrasts between the No Adverts, Irrelevant Adverts and Disreputable Adverts showed that credibility was significantly affected by both Irrelevant Adverts $F(1,120) = 34.29, p <.001$ and again by Disreputable Adverts $F(1,120) = 11.72, p <.05$. as seen in Table 13.

Table 13: Within Subjects Contrasts

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advertisements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Level 1 vs later)</td>
<td>31.449</td>
<td>1</td>
<td>31.449</td>
<td>34.289</td>
<td>.000</td>
</tr>
<tr>
<td>(Level 2 vs Level 3)</td>
<td>13.187</td>
<td>1</td>
<td>13.187</td>
<td>11.717</td>
<td>.001</td>
</tr>
<tr>
<td>Advertisements Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Level 1 vs later)</td>
<td>110.060</td>
<td>120</td>
<td>.917</td>
<td>.917</td>
<td></td>
</tr>
<tr>
<td>(Level 2 vs Level 3)</td>
<td>135.054</td>
<td>120</td>
<td>1.125</td>
<td>1.125</td>
<td></td>
</tr>
</tbody>
</table>

This significant difference in response to the Irrelevant Adverts vs the Disreputable Adverts supported H3.

Interestingly then, unlike the first between groups study, which only supported H2, this within subjects study supported H1, H2 and H3, and the directionality of the effects were consistent with expectations for both Appearance and Advertising on Credibility as seen in Illustration 34.
Illustration 34: Profile Plot of Appearances * Advertisements
7 Qualitative Analysis

7.1 Nature of Analysis
Some of the most interesting data from this research came in the form of comments that people made about the web pages they viewed. These comments shine extra light on the perceptions participants formed. This sort of Qualitative analysis was used by a number of studies considered in the literature review, including Fogg et al (2003), Alsudani & Casey (2009), Michailidou et al (2008), Hartmann et al (2007). Of these Fogg et al (2003) was most influential and most comprehensive in their approach.

7.2 Coding of Comments
Participants were offering comments prompted by the following question:

What is your overall assessment or opinion of this service?

This question was placed at the bottom of the survey, allowing participants an opportunity to formulate their thoughts and opinions as they answered the previous eight questions.

Once the studies were complete, the researcher examined the database of comments created, marking cases that showed evidence of someone submitting meaningless, essentially blank or repetitive comments. After marking these cases as void, the researcher went through the large volume of comments from both studies and assigned theme codes to identify themes and terms found in the comment.

Sparse or inadequate answers in the open ended comments were always a possibility in an anonymous online survey, as there is no interviewer guidance or encouragement available to draw out fuller answers, but overall, there were very few void or meaningless comments, of the 2854 comments captured, only 58 were marked as void, less than 2%. This may have been because the form could not be submitted without something being entered in the comments field, but a surprising number of participants seemed to have things they wanted to say.

Each comment often received more than one theme code, and any code could be positive, negative or neutral. For example, the comment below was coded, negatively, in three categories: Aesthetic, Credibility and Meaning

“Too gushy and seemed unbelievable. Yuck ads-very off putting.”
The set of theme codes, intended to reflect aspects of the overall issues being considered in this research and discussed in the literature review, is outlined in Table 14, with a very non-exhaustive example of concepts associated with each. Values given could be positive or negative.

**Table 14: Themes used in coding**

<table>
<thead>
<tr>
<th>Theme</th>
<th>Concepts addressed</th>
</tr>
</thead>
</table>
| Aesthetic | Comments to do with the appearance of the design, layout, images or advertising  
  e.g beauty, ugly, clear, clutter, confused, simple, complex, colourful, bland, attractive, unattractive, consistent, inconsistent, bright, dull, tidy, messy, distracting |
| Affect   | Comments reflecting the attitude or mood of the User  
  e.g good, bad, happy, sad, hate, love, annoy, delight, like, dislike, disgust, please, excellent, worthless, irritating, inviting, uninviting |
| Credibility | Comments about trust, honesty, professionalism  
  e.g credible, deceptive, honesty, dishonesty, trust, distrust, amateur, professional |
| Reputation | Comments about user reviews or perceptions of successfullness and experience  
  e.g reputable, disreputable, experienced, inexperienced, quality, shoddy, successful, unsuccessful, long standing, new |
| Meaning  | Comments about the information quantity, quality, relevance or appropriateness  
  e.g comprehensible, informative, uninformative, relevant, irrelevant, pointless, insightful, understand, misunderstand, interesting, boring |
| Interaction | Comments about the usability or legibility of the page  
  e.g. ease, difficulty, usable, unusable, legible, illegible, frustrating |
| General  | Any other comments |

After coding all the comments, frequencies for each code category were established, both for each survey and in total, in regard to positive, negative and neutral comments.

These frequency scores, seen in Table 15, give an indication of what issues people felt they were reacting to when they made their credibility evaluations in the previous 8 semantic differential questions.

The individual comments also contained insights into specific details and aspects that the quantitative research could not cover. This large (2854) collection of comments about Directory Website Credibility offers numerous opportunities for analysis, and some of these issues will be addressed in the discussion.
7.3 Coding Results

7.3.1 Overall Results

Overall, 41% of codes from comments were negative, 15.3% were neutral, and 43.7% were positive.

For individual themes across both surveys, the results were ordered as follows in Table 15:

<table>
<thead>
<tr>
<th>Theme</th>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaning</td>
<td>728</td>
<td>111</td>
<td>759</td>
<td>1598</td>
</tr>
<tr>
<td>Credibility</td>
<td>645</td>
<td>236</td>
<td>439</td>
<td>1320</td>
</tr>
<tr>
<td>Reputation</td>
<td>516</td>
<td>82</td>
<td>579</td>
<td>1177</td>
</tr>
<tr>
<td>Affect</td>
<td>495</td>
<td>168</td>
<td>442</td>
<td>1105</td>
</tr>
<tr>
<td>Aesthetic</td>
<td>259</td>
<td>26</td>
<td>428</td>
<td>713</td>
</tr>
<tr>
<td>General</td>
<td>52</td>
<td>55</td>
<td>30</td>
<td>137</td>
</tr>
<tr>
<td>Interaction</td>
<td>12</td>
<td>0</td>
<td>21</td>
<td>33</td>
</tr>
</tbody>
</table>

The largest number of comments made referred to Meaning related issues – about the information, its quantity and relevance. Importantly, this includes comments about the irrelevance or relevance of the advertising. There were slightly more negative than positive comments at 47.5%.

Credibility and Reputation came next, with, surprisingly, a positive bias for credibility and a negative bias for reputation, although this partly might be attributed to the high neutral result for credibility, at 17.9%.

Affect was also positive, although again with a high neutral, but Aesthetic was resoundingly negative, with a 60% result.

The other two themes, General and Interaction, made up less than 5% of comments.
7.3.2 Results for Between Groups study

For the first between groups study, 37.5% of codes from comments were negative, 15.9% were neutral, and 46.6% were positive.

For individual themes the results were ordered as follows in Table 16:

<table>
<thead>
<tr>
<th>Theme</th>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaning</td>
<td>578</td>
<td>91</td>
<td>522</td>
<td>1191</td>
</tr>
<tr>
<td>Credibility</td>
<td>510</td>
<td>174</td>
<td>277</td>
<td>961</td>
</tr>
<tr>
<td>Reputation</td>
<td>433</td>
<td>71</td>
<td>449</td>
<td>953</td>
</tr>
<tr>
<td>Affect</td>
<td>379</td>
<td>106</td>
<td>287</td>
<td>772</td>
</tr>
<tr>
<td>Aesthetic</td>
<td>199</td>
<td>15</td>
<td>225</td>
<td>439</td>
</tr>
<tr>
<td>General</td>
<td>52</td>
<td>55</td>
<td>28</td>
<td>135</td>
</tr>
<tr>
<td>Interaction</td>
<td>5</td>
<td>0</td>
<td>9</td>
<td>14</td>
</tr>
</tbody>
</table>

Again, the relative number of different types of issue raised and the positive and negative skew are very similar to the overall, with the exception of Meaning now being positive and Credibility being even higher. This increase in positivity is consistent to an extent with the quantitative results for the between groups study, as the results for that were in less significant overall than those in the within subjects study and at times trended in the reverse direction to expectations.

7.3.3 Results for Within Subjects study

For the second within subjects study, we see some major differences. 51.5% of codes from comments were negative, 13.4% were neutral, and only 35.1% were positive. Although having a smaller number of raw comments at 701, there were still 1796 individual comment codes recorded.

For individual themes the results were ordered as follows in Table 17:
Table 17: Within Subjects coding results

<table>
<thead>
<tr>
<th>Theme</th>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aesthetic</td>
<td>60</td>
<td>11</td>
<td>203</td>
<td>452</td>
</tr>
<tr>
<td></td>
<td>21.9%</td>
<td>4.0%</td>
<td>74.1%</td>
<td></td>
</tr>
<tr>
<td>Meaning</td>
<td>150</td>
<td>20</td>
<td>237</td>
<td>407</td>
</tr>
<tr>
<td></td>
<td>36.9%</td>
<td>4.9%</td>
<td>58.2%</td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td>135</td>
<td>62</td>
<td>162</td>
<td>359</td>
</tr>
<tr>
<td></td>
<td>37.6%</td>
<td>8.5%</td>
<td>45.1%</td>
<td></td>
</tr>
<tr>
<td>Affect</td>
<td>116</td>
<td>62</td>
<td>155</td>
<td>333</td>
</tr>
<tr>
<td></td>
<td>34.8%</td>
<td>18.6%</td>
<td>46.5%</td>
<td></td>
</tr>
<tr>
<td>Reputation</td>
<td>83</td>
<td>11</td>
<td>130</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>37.1%</td>
<td>4.9%</td>
<td>58.0%</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>7</td>
<td>0</td>
<td>912</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>36.8%</td>
<td>.0%</td>
<td>63.2%</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>.0%</td>
<td>.0%</td>
<td>100.0%</td>
<td></td>
</tr>
</tbody>
</table>

Now all the results are skewed to the negative, some of them with much larger proportions than any theme in the first study. The key change here is the sudden rise of Aesthetic, and that with a high negative majority of 74.1%

This change is dramatic, and is partly a result of moving to a within subjects design where all the participants can suddenly see the different designs and advertisement combinations in close proximity, and that comparative effect, inherently missing from the between groups design, has presumably begun to influence their opinions.

The effect of different combinations of Appearance and Advertising

To better understand what was happening, it is necessary to examine the comments made when participants were exposed to the different combinations of independent variables - Appearance and Advertising. Illustration 35 shows some interesting effects on some key themes:
It seems clear that while both appearances engendered negative Aesthetic comments, the Unattractive design had many more (44.5% vs 29.6%), and also far fewer positive comments (5.5% vs 16.4%)

An even more dramatic effect was seen in the impact of different types of advertising, with the Disreputable Ads scoring 34.7% and negative comments making up 74.1% of all Aesthetic comments about advertising.

Some example comments encountered were:

“The bright blue, green, Comic sans font, etc tell me that they don’t seem to mind if they aren’t well presented. They have no logo, and a stock photo of a smiling accountant, they offer an affordable service, but it looks unorganized.” Male, 25 -34yrs

“Ha, we seem to have gone backwards here. The flashing adds, gimmicky adds certainly detract from the seriousness of a small business provider. Also, tacky graphics make it look amateurish again.” Female 35 - 44yrs

“Very brief - shows lack of preparation which gives a shonky feel to it. The ads at top and bottom intrude and make the thing look very mickey mouse” Male, 45 - 54yrs

“Site looks cheap so Company looks cheap too.” Male, 35 - 44yrs

Similar but less dramatic results were seen for the effect of Appearance and Advertising on Meaning:

Illustration 36: Results for incidence and direction of Meaning Comments
While the result for appearance is not as clearcut and dramatic as the previous example, the effect of Advertising on Meaning comments discussing relevance and appropriateness is clearly seen, with the Disreputable Ads taking 43.0% and Irrelevant Ads taking 34.6% of all negative comments about Meaning, with a combined share of 77.6%.

Interestingly, many found the spelling and grammatical errors off-putting as well, and these were genuine, the text was exactly as sourced from other real online directories.

Some example comments encountered were:

“The spelling mistakes and poor quality tell you that the business is incompetent and unprofessional” Male 18-24yrs

“I probably would have navigated away from this page, hard to see the listing or think it will be reliable with such surrounding content. Seems the site takes the information from elsewhere, but it isn’t monitored.” Female, 25-34yrs

“Dreadful site with shonky grammar and spelling. The whole concept seems flawed. The obviously fake testimonials really put you off. Very dishonest looking. You can just imagine getting serious food poisoning here!” Male 45-54yrs

“The description of the service and the surrounding listing material were at odds with each other. It’s hard to see that location for your listing as something coincidental.” Female, 25-34yrs

Finally, the effect of the two independent variables on Credibility can be seen in Illustration 37:

Illustration 37: Results for incidence and direction of Meaning Comments
The result for appearance supports the overall results from the Quantitative study, although there is a noticeable trend for greater uncertainty reflected in the far higher proportion of neutral comments on this theme than the previous two.

The advertising again makes a major negative impression on perceived credibility, with with the Disreputable Ads taking 45.1% and Irrelevant Ads taking 32.7% of all negative comments about Credibility, with a combined share of 77.8%.

Some example comments encountered were:

“That Russian wife ad destroys any chance this site has of credibility. The rest is just a mish mash and thoroughly unconvincing. You can’t read the second testimonial with caps - yuk!” Male 45 - 54yrs

“Amateur and lacking professional input into the advertisement, therefore the service being offered.” Male 65+

“The advert above the company picture and map was very distracting and unsuitable for this type of service therefore made the information about the services seem untrustworthy.” Female, 35 - 44yrs

“Would have to be a dodgy accountant to advertise here.” Male 35 -44yrs
Discussion

8.1 Introduction

This section outlines a general discussion of the results presented in the previous section. It starts with a summary of the concepts, objectives, methodology and findings reported in this research. The effectiveness of the research and the contribution of this work to the body of knowledge will be discussed in Section 8.3.

8.2 Recapitulation

To help understand how the results fit into the framework of this research, a brief summary of the concepts, themes and objectives addressed so far is in order.

Section 1 introduced issues around user emotion, or Affect, outlined the history of web design and how affect had largely been unaddressed in traditional HCI studies until recently, and posed some questions about how exactly users are being affected by the content and designs they interact with.

Section 2 reviewed the literature on user affect, defining it (Norman, 2005) and followed that with an examination of various models of Aesthetics (Porat & Tractinsky, 2008), (Zettl, 1973) also exploring the particular impact of Aesthetics on affect as described by Lavie & Tractinsky (2004). Then the role of Advertising on user affect was considered by examining studies on the intrusiveness and irrelevance of advertising(Yoo & K. Kim, 2005), (E. J. Newman et al., 2004).

Various aspects and models of online trust and credibility were compared (Egger, 2000), (Riegelsberger et al., 2005), (Wang & Emurian, 2005), some parameters and components of credibility defined (Corritore et al., 2003) and mechanisms for the transmission of trust in systems discussed (Warkentin et al., 2010), (Stewart, 2003). The roles Aesthetics and Advertising might have on the formation of perceptions of trust and credibility was considered. (Karvonen, 2000), (Riegelsberger et al., 2004), (Ellison et al., 2006)

These sections were integrated into the context of a common online business model, the Multi Sided Online Market model as used by business directories, which extends the Two Sided Market Model (Rochet & Tirole, 2003). A new integrated model was formulated that mapped the Network of Affect in a Multi Sided Online Market such as an online business directory, and
three hypotheses were proposed to assess how Aesthetics and Advertising might impact users impressions about the credibility of services seen within such a hypothetical directory.

Section 3 outlined a research methodology for exploring these hypotheses, using a fictitious business directory website and online survey forms to assess credibility both from a quantitative and a qualitative perspective. Aesthetic design principles derived from the literature review were employed to create Attractive and Unattractive versions of this website, and example advertising selected to populate versions with Advertising.

Section 4 details how the final research instrument was tested using a small pilot study, refined, and then used in two studies, one between groups and the other within subjects.

Sections 5, 6 describe the process of participation in the two studies, then tables the results of each, assesses the validity of the research questions and the data, and performs various analysis to determine if the hypotheses are supported, finding that one (H2) was supported in the first study and all three in the second.

These findings were followed in Section 7, the previous section, by a qualitative assessment of participant comments about the website appearance, advertising and the credibility of the services seen.

### 8.3 Contribution

The hypotheses and objectives set at the end of Section 2 were:

- **H1.** Reduced website attractiveness negatively affects users perceptions of the credibility of the services represented by the directory.

- **H2.** The presence of website advertising of low relevance negatively affects users perceptions of the services represented by the directory.

- **H3.** The presence of website advertising of low relevance and low reputability negatively affects users perceptions of the services represented by the directory more than low relevance alone.

Additionally, emergent issues around the role of reputation, aesthetics, user affect and advertising on the users perceptions of all parties in this Multi Sided Online Market were to be explored in a qualitative component.
How fully these objectives have been attained will be discussed by reviewing the types of knowledge uncovered in this research.

The quantitative and qualitative analysis of these two studies produced a set of conclusions and trends that can help expand the understanding about the relationship between Aesthetic elements, the Presence of Advertising and how these things affect users’ perception of the credibility of a service seen on a web page. It adds robust information, building on previous studies on how users perceive the visual appearance of a web page (Tractinsky & Lowengart, 2007) (Hartmann et al., 2007), and how they react to Advertising (Cho & Cheon, 2004) with specific regard to how these things impact their affect, and how that then modified their impressions of credibility.

This issue of the impact of affect on credibility of a website had been explored, but only within the context of normal consumer to business websites like e-commerce sites, a business model where the website represents the business directly, as a storefront or brochure. This study extended this previous work by exploring how those same forces impact all the participants of a Multi Sided Online Market, like an online business directory, where the website is a mediator between parties not a representative, and both users, businesses with listings and even third party advertisers all share access to the one platform and combine to shape the nature of the content and presentation it exhibits. The three hypotheses above were developed to explore some of these effects.

All three hypotheses statements, H1, H2 and H3 are supported from the final study, but only one, H2 from the first study.

H1 stated that reduced website attractiveness negatively affects users perceptions of the credibility of the services represented by the directory. This was not supported by the first between groups study, where each individual only saw one version of the website design. In fact the trend, while not significant, was in the opposite direction, with a small positive effect seen for the Unattractive Design. However, the second study, a within subjects design with inherently more statistical power, and where all the website designs were seen by the same people, found significant support for this hypothesis.

This conclusion supports the idea that the Aesthetic appearance of a page in a business representing a Multi Sided Online Market, such as an online directory website, will have an impact on how the service depicted on that directory will be perceived in terms of credibility, extending the current understanding in the body of knowledge regarding this effect with ordinary websites. This quantitative result was also strongly supported by the qualitative
analysis of the participant comments. Some potential business and social implications of this, and thoughts for further research, are discussed later.

This result also supports the value of the Design Principles gleaned from the literature that were used in the production of both the Attractive and Unattractive designs, but it does raise some issues around best methodologies to use for Aesthetic studies. Both of these issues are also explored in more detail later in this discussion.

The second hypothesis, $H_2$, states that the presence of website advertising of low relevance negatively affects users perceptions of the services represented by the directory. Both studies significantly supported this hypothesis. Importantly this means the impact of this effect has now been measured both between groups and within subjects, and in both cases was found significant, implying that the negative credibility impact of irrelevant advertising is a somewhat absolute effect based on its mere presence rather than an artefact of comparisons made by individuals exposed to pages with Advertising and without Advertising in the same experiment. The business and social issues pertinent to this are discussed below.

Again the qualitative results confirmed this finding, and raised interesting insights into the strategies and cascading sets of criteria users depend on to formulate their judgements about the things they are seeing when hard information is unavailable. This whole issue of how users form such judgements online, such as the use of Warranting (Warkentin et al., 2010), which was discussed in the literature review, is revisited below in the light of these comments.

$H_3$ states that the presence of website advertising of low relevance and low reputability negatively affects users perceptions of the services represented by the directory more than low relevance alone. This effect was only able to be assessed in the second study, but the results again gave significant support to this hypothesis. The Disreputable Advertising was also strongly identified as a negative, inappropriate and credibility reducing feature of the qualitative analysis, with many comments directly identifying the offending ad, the website platform itself and the service as discreet but connected items with comments like:

"I would immediately leave this website because of the nature of the advertising. I would not trust that the information would be accurate because the main focus of the site appears to be the advertising not the information." Female, 45 - 54yrs

"For a professional this site was very poor and the company did not look good in any way. I dislike all these sites as they are covered in other adverts which detract from what you want to see. I would not trust my money to this company." Male, 45 - 54yrs
This ability of users to clearly discern distinct social and relational components that made up the content and context of the page they were viewing, lends weight to the Multi Sided Network of Affect model proposed in the literature review, where the various sides of the Market, the users, businesses, advertisers and the platform itself function as something of a neighbourhood, a mutually understood social context within which certain types of association are considered beneficial, and others are considered negative. It provides evidence that users were not simply conflating the services and the advertising and the platform itself into one amorphous “site” experience, but instead saw each part for itself, distinct yet still inextricably interrelated by their shared context and presentation.

8.3.1 Implications for Multi-Sided Online Platforms

The Multi Sided Network of Affect Model proposed in the literature review identified four potential interested parties when considering the dynamics of an online business directory. At the centre is the Platform, the directory website itself, often a business that exists simply for that purpose and has no other source of funds, unlike the monolithic telecommunications organisations that historically created printed directories.

On three sides of this directory are the business services listed on the directory, the users of the directory who are seeking the services listed, and advertisers who provide a source of income for the directory, often effectively subsiding what is otherwise a free service for both users and often the businesses listing as well. Each party in this group can play a role in the content and appearance of the directory.

Implications of Aesthetic Design

At the most fundamental level, it is the platform itself that determines the overall aesthetic of the directory. And good design, while not universally appreciated, does universally cost money; expenditure which may not be justified without a full understanding of the consequences of a given strategy on a given audience. Depending on the type of services they list and the type of customers they target, some directories may require substantial investment in Aesthetic design. The exact mechanics of determining this are not clear, but this study has shown that whether or not the platform owners understand the issue, their users will be reacting to the design of their site, or lack thereof, often by leaving altogether. Comments from users of the Unattractive site like:

“\textit{You cannot have a credible advert for an elegant dining restaurant under an advert for dating hot russian girls. it puts me off the idea of eating.}” Female, 18 - 24yrs
“The webpage appears outdated and tacky. The irrelevant advertising would put me off, and overall the layout and the colour theme used seems unprofessional.” Male, 18-24yrs

“I would immediately leave the site upon opening it.” Female, 45-54yrs

“Little care and effort went into preparing this, and you can tell. It looks like they don’t really want your service, so I wouldn’t bother going” Female, 18-24yrs

imply a high sensitivity of users to the overall presentation of the website, and a lack of inherent confidence in websites they deem cluttered, ugly or amateur.

Previous studies have already shown clearly that there was a significant negative effect on Website credibility of unattractive designs (J. Kim & Moon, 1998), (Karvonen, 2000), and the quantitative results from this research confirm this, but the comments here speak to the social context of this unease, users often seem to be visualising not only the kinds of services that might allow themselves to be seen on such a site, but also seem to be judging the kind of people that might use it, and then determining if they fit in, or should leave the neighbourhood. Comments like those above suggest users clearly see the design as reflecting an overall measure of care, and mistrust even small negative indicators as indicative of a larger trend, rather like the famous, if apocryphal, “Broken Window Effect” seen in physical neighbourhoods (Wilson & Kelling, 1982).

This effect is likely to impact on usage of the website, which, due to network effects seen in multi-sided platforms, will tend to reduce the number of business listings, and also the amount and quality of advertising available as an income stream for the platform. So being ugly is likely to be a bad thing for the growth and viability of the Platform itself.

Businesses who choose to list on directories will also be caught up in the negative effect of bad design. The quantitative results of this research, in the second study, show clearly that there was a significant negative effect on services credibility from the Unattractive Design. Many comments also showed that the users were often well aware the website and the service were distinct, and in consequence they judged the services for allowing themselves to be seen on such a platform.

Implications of advertising relevance

The relevance of advertising displayed is partly a choice of the Platform, and partly under the control of the Advertiser via functionality like Adsense which detects keywords in the page and tries to match ads to the apparent subject. But the advertisers choices on ads are, as discussed in
the literature review, in a large part determined by the number of users and perceived quality of the site (S. P. Anderson & Gabszewicz, 2006). If profits are largely driven by advertising, this can tend towards a negative feedback loop (Beuscart & Mellet, 2008) where worse ads appear to fewer users and compete with attention on a directory with fewer services, and so on in a downward spiral.

But adverts, unlike aesthetics, can be inherently not credible or untrustworthy due to their content as well as their appearance and in both the between groups and within subjects studies, the adverts had a significant negative effect on the perceived credibility of the services. The greater the lack of credibility of the advert perceived by the user, the deeper was the effect on the services, as established by fact that the second study saw even stronger negative effects for disreputable ads than for merely irrelevant ones.

While the ads were often noted as bringing negative effects aesthetically by introducing distracting animation, different fonts, colours, and styles that undermined any brand values and generally added clutter to the website, the users saved their most dismissive comments for appropriateness of message and social meaning of the most disreputable ads:

“When a site is accompanied by adverts for casinos and Russian brides it means it is not up to much” Male, 45-54yrs

“Trashy adverts all over the place smack of cheapness.” Female, 35-44yrs

“Surely they wouldn’t know that hot Russian girls were being advertised on their website? As accountants that is highly inappropriate and unprofessional.” Female 35-44yrs

Neither the businesses, nor the users would seem to gain any benefit from the presence of irrelevant or disreputable advertising on a directory. The negative view of the services engendered can only reduce the number of users, and one would assume the number of external transactions actually carried out. There is an argument to be made that even low grade advertising can in fact be seen as a necessary evil, by the simple virtue that its presence helps subsidise the existence of these directory platforms with no cost to the users and often no cost to the services, increasing the overall 'good' to society at least in principle (S. P. Anderson & Gabszewicz, 2006). But in practice with multiple directories competing for a fixed pool of users and services, the real amount of 'good' created by subsidising numerous low credibility directories seems unclear.
8.3.2 Implications for Affect & Credibility Research

This research confirms the results of previous study of user affect on credibility but extends this knowledge into a more complex, layered model of human transactions and expectations.

The significant results of the quantitative studies demonstrate that perceptions of credibility are indeed transferred from a host website to the services hosted, and many of the comments also give support to the idea that the inherent untrustworthiness of the advertising was directly conflated in the minds of the users with the trustworthiness of the services.

This result shows that credibility is transferrable and mutable as suggested by Stewart (2003), not discrete or contained to a trigger or cause. Instead, negative credibility seemed to be contagious, and once identified, it proved capable of tainting everything it touched.

Reasons for the extremity of this behaviour are clear from the literature (S. Tseng & Fogg, 1999), (Chopra & Wallace, 2003) and essentially reduce down to the tendency for people to interpret interactions with systems or even inanimate objects as if they were dealing with another person, giving each facet of the transaction the weight of a social actor.

The natural tendency to interpret and assess other people we interact with from social and physical cues (Norman, 2005) is mirrored in the way users have judged the services in this research. Users indeed actively seek confirmatory information, if solid information is missing, they search for clues wherever they can find them (Ellison et al., 2006), demonstrating a powerful bias to swift judgement. Of the possible cues available for their assessments in this research, users found Advertising the most compelling, rather than Appearance. Why this might be is an open question.

Perhaps advertising is understood as a more tangible form of social actor than appearance; advertising is after all opinionated, by nature it is trying to communicate a message, it is more like a ‘voice’ or a person than the appearance of the site, which more closely resembles a setting or an environment, a stage for the actors as it were. More research is needed to explore this issue, but it seems clear from the comments and the quantitative results that users had a greater emotional response to advertising they didn't like than to designs they didn't like.

8.3.3 Implications for Web Design Practice

According to Alberti, considered one of the most influential of Renaissance architects, “it is the property and business of the design to appoint to the edifice and all its parts their proper places, determinate number, just proportion and beautiful order” (Eden, 1943).
This elegant classical vision of the nature of Design is compelling, and seems somehow self evident, but in practice there is little agreement on the actual rules and boundaries of this 'just proportion and beautiful order', or how disparate and essentially random elements might be coerced into such a state of harmony.

Numerous efforts to tame this strangely unruly sense of order with fixed principles have been made down through history. From the Greek Golden Mean, Leonardo's Vitruvian Man, and Alberti's aforementioned “De Re Aedificatoria” right up to the twentieth century’s “Swiss” or International design styles, the great and the small alike have all attempted to apply rules, lines and measures to understand design. Even the humble CSS Grid based designs of the modern web are descended from principles and technologies that Gutenberg was employing when he created the printing press.

So in the spirit of these various recurring attempts, this research drew on a number of coexisting sets of design principles (Everard & Galletta, 2006), (Michailidou et al., 2008), (Sicilia & Ruiz, 2007), (Zettl, 1973) as discussed in the literature review, and formulated one of many possible sets of Design Principles, a concrete set of Aesthetic components, that while still subjective, at least provided a framework for testing, and then proceeded to put those Principles to the test.

While necessarily incomplete, and one of many similar models, the results of the studies show that the Aesthetics of design can be grounded, successful designs can be defined to some extent empirically, and unsuccessful ones also. It's perhaps possible to say that a design is good, to some extent, or bad, and have a reason beyond personal taste. The real issue however, is whether this matters to anyone other than designers.

Does the audience of the web need to care about the design, or is it enough just that it affects them. Unhappily, even if the audience can be shown to care, it is often not they, the audience of users, for whom the site is actually being designed. It is the interests of the businesses and organisations the websites are representing that are usually being served. And in this environment, financial concerns usually override vague and amorphous concepts like Design, and financial concerns are often best served by making sure that any possible 'whitespace' in the design is filled with the maximum number of paid advertisements.

But then, with regard to the visual impact of advertising, it has already been shown that users do indeed care, and this study enlarges on that understanding by examining their responses both in a more realistic, isolated, between groups context and an unrealistic but effective within subjects comparative context. And in both formats users generally hated, and were distracted, and were disgusted, by the advertisements. The Advertisements themselves were ugly, but

Discussion
importantly, they were seen to make the website ugly, and the content untrustworthy. And many users made it clear they would not stay.

Unlike a physical marketplace, the whole internet is always just one click on the back button away, and if users don't like your site, they will simply leave it. This is perhaps not news to any user of the internet, or to any web designer, but this research has helped to again confirm that poor design and excessive advertising can degrade the meaning and experience of a website to all users, not just designers of rarified personal tastes, but actual customers.

This information has value then, to those who employ the designers to build the sites. Design has the power to hold users attention, and therefore potentially add value to the bottom line. Excessive advertising has the power to make users leave. These are powerful arguments for good design and minimal advertising even if no intrinsic value for these things can be seen.
9 Overall Limitations & Future Research

9.1 Limitations

Clearly, this study is subject to some limitations. The research was short in duration, small in scope, had low resource requirements and was therefore achievable by a single researcher. But since the data gathered needed to remain anonymous for the sake of ethical considerations, and the actual physical process of website interface usage, user affect, physical motions and order of events and behaviours could not be observed but was instead assessed through an anonymous online survey, the full richness of user affect cannot be determined.

Similarly, in principle, since both surveys were anonymous, the results might have been influenced by any subjects that attempted both surveys or one survey twice. In that case familiarity with the web site or the services could have influenced the users’ impressions and evaluations, their answers may have been tainted by their past interactions with the surveys.

In practice however, care was taken to ensure that these anonymous respondents did not in actual fact do any survey more than once. This was achieved firstly by use of a cookie set in their web browser that retained information about previous completions of the survey and did not permit users to begin a new session from the same browser and machine. While this measure could be circumvented, the main means of assuring subject independence was that any participant wanting a movie ticket for their compensation had to enter a physical address, and often an email address, to permit the ticket to be sent. While this data was not stored as part of the survey data for ethical reasons, nonetheless an independent party charged with monitoring these addresses and mailing out vouchers was able to confirm that no duplicate addresses were provided. So while not impossible, it would appear that the likelihood of individuals attempting either survey twice is very low.

Limitations of the First Research Design

As discussed in the results section for the between groups study, there was a weakness in the design of the first study, in that the presentation of services was left to the participants discretion, and was not counterbalanced. This lack of counterbalancing meant that examining the effects on individual services or different types of advertising was not feasible. The design was however still perfectly capable of addressing the first two hypotheses, and \textbf{H2} was supported by the results and analysis.
This lead to the decision to run a second study, using a within subjects design, which has inherently more statistical power, and this study confirmed H1, H2 and H3.

The lack of a significant result for the first study with regard to aesthetics, even though the website designs were identical, could merely be an artefact of the reduced power of between groups designs, or it might indicate that exposing participants to all variations of the design in the second repeated measures study increased their sensitivity to the visual differences. It is possible for instance, that if the repeated measures of the second experiment had used a larger time gap between measures rather than being sequential, the results might have more closely resembled the between groups results for aesthetics. It's an open issue that definitely needs more research.

**Moderating effect of user reviews**

Although the study could not feasibly explore the effect of reputation through user reviews as directly as might have been desirable, a mix of real world user reviews were added to the services, largely to increase the plausibility of the content. This could, on first examination, have introduced some confounding variables into the assessment of credibility.

However the quantitative aspects of the study were unaffected in practice, as care was taken to ensure a balanced mix of reviews for each type of service presented, and since the results measured in both studies were assessed from the combined data from across the entire set of services, individual service variations were not featured.

Some comments did mention the absence, presence or content of reviews both positively and negatively for specific services, and these tended to reinforce the already known findings about how users interpret and depend on other users feedback (Toms & Taves, 2004) discussed in the literature review, but the comments also mentioned similar service dependant artefacts like the quality of individual photos, grammatical errors in the descriptions and even the names of the services themselves. This data effectively became part of the unique variations in aspects of each service, similar to the photos, description details, spelling mistakes and other realistic variance that came from the original source material.

Ideally, future research would facilitate the addition of reviews as a true independent variable, but that was outside the scope of this study.
9.2 Future Research

Many aspects of this research could be improved in future efforts. More work refining the Design Principles used in this study into a generalisable form would be of benefit, and other attempts in this direction (Michailidou et al., 2008), (Ivory & Megraw, 2005) that have contributed to the body of understanding about Aesthetics could be integrated into a more robust model of exactly what defines attractive versus unattractive. Generalising these design principles into a broader testing framework is perhaps a difficult task, for no set of rules seems able to fully represent the wealth of possibility of the human creative spirit. But developing more complete and reproducible measures is a prerequisite for any future study of the impact of the aesthetic aspect of web sites on user affect and credibility.

There is also room to consider the aesthetics issue from the reverse position, to examine what excellence of design might bring positively to the perceptions of services seen on a directory. Can excellent design actually enhance a weak service? Similarly the advertising issue could be considered from a positive perspective, is it possible to produce a website with advertising so well integrated, so relevant and so pleasing that it actually enhances peoples perceptions of the services it surrounds. That would be a fascinating pursuit indeed.

There is definitely a need for a more conclusive comparison between the two research methods employed, to ascertain if users are responding more strongly to aesthetic differences seen in the repeated measures design partly because the differences are simply being made more obvious, and if so some consideration of how this impacts the validity of this approach for aesthetic research needs to be undertaken.

It would also be interesting for future research to focus on aspects that were neglected in this project due to lack of resources, such as better understanding the role of user reviews in moderating affect and the impact of all these effects on individual types of services with different inherent credibility.

Additionally, research that examined some of the other dimensions of this Multi Sided Network of Affect would be of benefit, as this study was only able to assess the users opinions of the services, whereas the model supports the users view of other users, and the businesses view of other businesses and of users. Even the views of the third party advertisers would shed light on the true dynamics of this model. These studies would require significantly more resources than were feasible for this research, but without this work, we are only glimpsing a part of the full richness of social interaction at work in the model.
10 Conclusion

This research was intended to reveal useful insights about currently poorly understood connections between a user's website interactions and their attitudes towards the content of the website, specifically the services and products that the website may represent or depict.

This issue has wider sociotechnical significance than merely determining the click-through rate of frustrated users on online directories, as the “products” on display can be people's livelihoods and the effects of the interaction can potentially reach beyond the boundaries of the screen into people's businesses and careers.

Used well, online mediators like the Multi Sided Markets studied in this research can facilitate real events more smoothly and economically than any real market or forum. The existence of eBay and Facebook are testaments to the power of unified human communication through social media. They are successful, thriving virtual neighbourhoods, sociotechnical success stories where real needs are being met to such an extent that people willingly choose to invest time in these virtual spaces at the expense of time spent in ‘real’ activities (Ellison, Steinfield, & Lampe, 2007). While they are not without their perils, they have the power to enrich life by successfully enabling what would otherwise be impossible or infeasible.

Online directories also show the potential to greatly enhance users access to services and increase the cost effectiveness for businesses promoting themselves online (Rysman, 2004). Unlike the aloof mechanisms of traditional business directories, which are handed out complete and fully formed for the consumption of the masses without room for their commentary, online directories can close the feedback loop by allowing users to comment and review services, and in the process paradoxically increase the value of the platform for both the users and the businesses seen on it, by increasing the pool of engaged users, enhancing their trust in the information and reducing the cost of information discovery (Jullien, 2005).

But when the systems and platforms and directories mediating the transactions fail to support the full needs, both practical and emotional, of those people who are visiting and transacting on their websites, they risk becoming at best abandoned ghettos on the web, or degenerate into lurid red light districts filled with flashing messages for services users really don't want. This is the social-technical gap at work, often perhaps caused by a fundamental misapprehension in the minds of the platform owner of the real purposes and meaning such systems have to the actual consumers and businesses who use it.

It is doubtful that the owners of the many low credibility online directories surveyed in the process of this research set out to provide anything other than a useful service for user and
businesses alike. Yet this study shows that it is all too easy to engender negative emotions by neglecting the human aspects of the mediating platform. Malice is not required, just a lack of vision. A sociotechnical system like a Multi Sided Online Market must be conceived and built as if it were an environment, a pleasant online neighbourhood where people might actively want to meet each other, and spend time. Ignoring human emotional responses in design is to repeat the same failure of the imagination that lead many post war modernist housing estates to swift decline and demolition. It is not enough to technically meet the needs of users with a ‘machine for transacting’ if the device itself is so intrusive that it inhibits the desire to transact at all.

Yet this pursuit of function at the expense of form is a recurring failure in technology design. The experience of interacting with a system invariably changes our attitudes and performance on that system. Tractinsky et al suggest that “positive affect is likely to improve decision making and creativity. Thus, the positive affect created by aesthetically appealing interfaces may be instrumental in improving users’ performance as well” (2000, p. 141).

Clearly, there is a need here to better understand the mechanisms at work, but despite attempts to quantify aesthetic quality with principles and models as seen in this study, and many before, we are not yet at the point of a Science of Aesthetic Design. Good design remains for now a product of human creative whimsy, frustratingly unpredictable and fugitive.

And even good online neighbourhoods can “go to the dogs”. While good design can help make an online directory a place where all parties can transact and communicate with confidence and clarity, these platforms often contain the seeds of their own destruction. With the advent of so many competing ‘freemium’ advertising-supported directories, the potential exists for the excessive or desperate pursuit of the advertising dollar to send these platforms into a downward spiral of perceived credibility. Strangely, it seems, at least from the findings of this study, that it is these third parties, Advertisers, fellow actors sharing the stage as it were with the users and businesses doing the actual transacting, who have the greatest impact on perceptions of credibility. The neighbourhood itself may be good or bad, but it is bad neighbours that seem to be the true problem.
11 References


References


Egger, F. N. (2000). Trust me, I'm an online vendor: towards a model of trust for e-commerce system design. In CHI’00 extended abstracts on Human factors in computing systems (p. 102).


Applications, 3(2), 176–189.


References 121


References


Appendix 1 - Design Source Websites

Example Website Designs

These detail, results and home pages from the following largely New Zealand based websites were used as examples for the aesthetic design of the research instruments.

http://esdot.co.nz
http://savebarn.co.nz
http://www.topmaq.co.nz
http://www.primeimports.co.nz
http://www.goldenlanterns.com
http://www.tamilbizcard.com
http://www.scratchandsave.co.nz
http://ancglobal.net
http://www.yalwa.co.nz
http://www.zipleaf.co.nz
http://www.nzdirectory.co.nz/
http://www.websitedirectory.co.nz/
http://www.hotfrog.co.nz
http://www.finda.co.nz
http://www.yellow.co.nz
http://www.yellow.com.au
http://www.nzsbdirectory.co.nz
http://www.shopnewzealand.co.nz
http://www.tradeboss.com
http://www.countrycreekconnection.com
http://www.looksie.com
http://go4newzealandbusiness.com
http://indiacatalog.com
Example Website Screenshots

countrycreekconnection.com

finda.co.nz

Appendix 1 - Design Source Websites
Appendix 1 - Design Source Websites
Appendix 1 - Design Source Websites

shopnewzealand.co.nz

tamilbizcard.com

shopnewzealand.co.nz

tamilbizcard.com
Appendix 1 - Design Source Websites
Appendix 1 - Design Source Websites
Appendix 2 - Pilot Study

Example Participant & Ranking Questionnaire


Gender: M / F

Internet Experience: None Limited Good Excellent

Website Ranking

1

2

3

4

5

6

7

8

9

Ad Ranking

1

2

3

4

5
Example Web Page Questionnaire

Questionnaire for Webpage Designs

Design: Finda

Circle the number under the terms that best describes your feeling about the webpage you've just seen

Overall Reactions to the Webpage

<table>
<thead>
<tr>
<th>Terrible</th>
<th>Wonderful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ugly</th>
<th>Beautiful</th>
</tr>
</thead>
<tbody>
<tr>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dull</th>
<th>Stimulating</th>
</tr>
</thead>
<tbody>
<tr>
<td>0  1  2  3  4  5  6  7  8  9</td>
<td></td>
</tr>
</tbody>
</table>

What was the most attractive feature?

What was the most unattractive feature?

Appendix 2 - Pilot Study 133
Appendix 3 - Information, Consent and Survey Forms

Information Sheet

Business Directory Website Research - START_DATE

INFORMATION SHEET

Researcher Introduction
Researcher: Daron Robinson
Research Type: 120 credit Student Research Project for completion of a Master of Information Systems.

Project Description and Invitation
I would like to invite participants for my Masters research project. This research explores the perceptions and opinions that participants form about companies and individuals that are represented on a fictitious business directory website. A series of short online questionnaires is completed by the participants, and the entire process can be completed online in the participants own time in under one hour.

Participant Identification and Recruitment
Participants have been recruited from within the Massey Albany Student body and elsewhere by word of mouth and by the use of Posters. However, the only criteria needed to participate is that the participant can use a web browser and is over the age of 18.

The number of participants will be in the range of 40 to 100. 40 is the minimum needed for the desired degree of accuracy in the research, and 100 is the maximum that funding will support. Each participant will be offered a movie voucher by way of compensation (or kaha) for their time.

Project Procedures
The participants will use a fictitious business directory website to examine information about companies and individuals represented on it and complete a brief online questionnaire for each one. This process is expected to take under one hour.

Data Management
The data from the participant questionnaires will be analysed as part of the researcher’s Masters research project, expected to be completed by February 2011.

The questionnaires data your provide may be stored for a further two years to facilitate any future publication of the research findings.

A summary webpage of the project findings will be available to each participant once the research project has been graded.

Your identity will remain entirely anonymous, any contact details you provide to receive your movie voucher will not be saved, or used in any way as part of the research.

Participant’s Rights
You are under no obligation to accept this invitation. If you decide to participate, you have the right to:
• Decline to answer any particular question
• Withdraw from the study by not completing the online questionnaires
• Ask any questions about the study at any time during participation
• Provide information on the understanding that your name will not be used unless you give permission to the researcher
• Be given access to a summary of the project findings when it is concluded.

Project Contacts
The researcher is Daron Robinson, and can be contacted by emailing daro@emaproteamedia.com or by calling 021 68 06 22. If you have any questions about the project, please feel free to contact the researcher:

Ethics Approval Statement
This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University’s Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O’Neill, Gowerster (Research Ethics), telephone 06 350 5240, or email humanethics@massey.ac.nz.

Download this Information Sheet as a pdf

Click here to participate in this research now
Consent Form

<table>
<thead>
<tr>
<th>Massey University</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Management and International Business</td>
</tr>
<tr>
<td>Massey University - Albany, Private Bag 112294</td>
</tr>
<tr>
<td>North Shore HSC, Auckland, New Zealand</td>
</tr>
</tbody>
</table>

Business Directory Website Research - January 2011

CONSENT FORM

Please note: ALL questions need to be answered for the form to be completed.

PARTICIPANT INFORMATION
The information provided in this form will be used in the research, but your identity will remain anonymous.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Please select...</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Internet Experience</th>
<th>Please select...</th>
</tr>
</thead>
</table>

PARTICIPANT CONSENT
I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

☐ I agree to participate in this study under the conditions set out in the Information Sheet.

Submit